

STIC-Biotech/ChemLib

191514

From: Chan, Christina
Sent: Wednesday, May 31, 2006 1:22 PM
To: Marvich, Maria; STIC-Biotech/ChemLib
Subject: RE: 09/729658 SEQ ID NO:2

~~Please rush.~~ Thanks Chris

Chris Chan
TC 1600 New Hire Training Coordinator and SPE 1644
(571)-272-0841.
Remsen, 3E89

CRFB

-----Original Message-----

From: Marvich, Maria
Sent: Wednesday, May 31, 2006 1:21 PM
To: Chan, Christina
Subject: FW: 09/729658 SEQ ID NO:2

29301

Please rush an after-final sequence search that is otherwise in condition for allowance.

Please include interference.

Thank you.

Maria Bonovich Marvich, PhD
United States Patent and Trademark Office
Remsen 2B84
AU 1633
Mail Box 2C70
571-272-0774

Searcher: Thibbe
Searcher Phone: _____
Date Searcher Picked up: _____
Date completed: 6/8/06
Searcher Prep Time: 5
Online Time: 5

Type of Search /
NA# _____ AA# _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure #: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable
STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: Compugen
WWW/Internet: _____
Other (Specify): _____

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: June 2, 2006, 17:21:28 ; Search time 200 Seconds
(without alignments)
893.858 Million cell updates/sec

Title: US-09-729-658C-2

Perfect score: 2098

Sequence: 1 MGYEVEVERELLPAAAPRR.....NMSKHTTFPGAIRLGEAPAS 391

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq.8.*

- 1: Geneseqp1980s.*
- 2: Geneseqp1990s.*
- 3: Geneseqp2000s.*
- 4: Geneseqp2001s.*
- 5: Geneseqp2002s.*
- 6: Geneseqp2003as.*
- 7: Geneseqp2003bs.*
- 8: Geneseqp2004s.*
- 9: Geneseqp2005s.*
- 10: Geneseqp2006s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2098	100.0	391	5	ABP53104 Human ect
2	2098	100.0	391	5	ABP53104 Human ect
3	2098	100.0	391	6	ABR42324 Human EDA
4	2098	100.0	391	7	ABO01928 Human EDA
5	2098	100.0	391	7	ADG35224 Human TNF
6	2098	100.0	391	7	ABW02286 Human EDA
7	2098	100.0	391	9	ADK06863 Cyclin-de
8	2098	100.0	391	9	AED55185 Ectodyspl
9	1965	93.7	391	9	AED55186 Murine ec
10	1873	89.3	377	5	ABO09083 Mouse Tab
11	1873	89.3	377	7	ABO01929 Mouse Ta
12	779	37.1	412	7	ADP75164 Fusion co
13	759	36.2	410	7	ADP75166 Fusion co
14	700	33.4	135	2	AAW44764 Human anh
15	294	14.0	357	2	AAW95115 Interveni
16	294	14.0	357	2	AAW57646 Collagen
17	294	14.0	357	4	AAW57646 Collagen
18	294	14.0	357	4	AAW57646 Collagen
19	286	13.6	742	8	ADL91066 Repetitiv
20	282	13.4	742	4	AAW57646 Collagen
21	282	13.4	742	8	ADL91057 Amino aci
22	282	13.4	742	8	ADL91057 Mouse col
23	280.5	13.4	1838	2	AAW53257 Human col

24	280.5	13.4	1838	7	ADP55566 Human Pro
25	280.5	13.4	1838	7	ADP55566 Human Pro
26	280.5	13.4	1838	7	ADP55566 Human Pro
27	280.5	13.4	1838	9	ADV70238 Tumor-ass
28	280.5	13.4	1838	9	ADZ09822 Human bre
29	280.5	13.4	1838	9	ADZ09822 Human bre
30	280.5	13.4	1838	9	ADZ09822 Human pro
31	277	13.2	330	2	AAW57645 Collagen
32	277	13.2	408	2	AAW07539 Collagen
33	277	13.2	408	4	AAW07539 Collagen
34	277	13.2	408	4	AAW07539 Collagen
35	276	13.2	622	8	ADL91065 Human col
36	276	13.2	622	8	ADP56200 Human PRO
37	275.5	13.1	180	9	AE001252 Wound dre
38	275.5	13.1	180	9	AE001252 Wound dre
39	275.5	13.1	180	9	AE001252 Wound dre
40	275.5	13.1	180	10	AEF82126 Cell cult
41	275	13.1	1466	4	AAE02534 Bovine al
42	275	13.1	1466	4	AAE02534 Bovine al
43	273.5	13.0	1014	9	ADW99577 Human gel
44	273	13.0	342	4	AAW39686 Human pol
45	273	13.0	547	3	AAW77985 Human col

ALIGNMENTS

RESULT 1
ABP53104
ID ABP53104 standard; protein; 391 AA.
XX
AC ABP53104;
XX
DT 11-NOV-2002 (first entry)
XX
DE Human ectodermal dysplasia protein (EDA) SEQ ID NO:6.
XX
KW Human; adipocyte complement related protein; zacrpx2; antiinflammatory;
KW vasotropic; antimicrobial; vulnary; antiarteriosclerotic; gene therapy;
KW adipogenesis; gluconeogenesis; lipogenesis; ischaemia;
KW glucose uptake; protein synthesis; thermogenesis; oxygen utilisation;
KW inflammation; neurotransmitter; antimicrobial; infection.
XX
OS Homo sapiens.
XX
PN WO200246417-A2.
XX
PD 13-JUN-2002.
XX
PF 07-DEC-2001; 2001WO-US047348.
XX
PR 07-DEC-2000; 2000US-0254019P.
XX
PA (ZYMO) ZYMOGENETICS INC.
XX
PI Haldeman BA, Thayer EC, Sheppard PO;
XX
DR WPI; 2002-657398/70.
XX
PT Novel human adipocyte complement related protein, zacrpx2, useful for
PT modulating energy balance in mammals, and preventing injury due to
PT ischemia/inflammation.
XX
PS Disclosure; Page 111-112; 118pp; English.
XX
CC The present invention describes a human adipocyte complement related
CC protein (I) designated zacrpx2. (I) has antiinflammatory, vasotropic,
CC antimicrobial, vulnary and antiarteriosclerotic activities, and can be
CC used in somatic cell gene therapy, and as a modulator of
CC neurotransmission. (I) is useful for modulating energy balance in mammals
CC or for protecting endothelial cells from injury, for modulating cellular
CC metabolic reactions e.g. adipogenesis, gluconeogenesis, glycogenolysis,
CC lipogenesis, glucose uptake, protein synthesis, thermogenesis and oxygen

PA (CHEN/) CHEN J.
PA (RAMA/) RAMANATHAN C S.
PA (XIAO/) XIAO H.
PA (GUAN/) GUAN B.
PA (BOWE/) BOWEN M A.
XX
PI Carroll PM, Chen J, Ramanathan CS, Xiao H, Guan B, Bowen MA;
XX WPI; 2005-756087/77.
DR
XX
XX New Drosophila tumor necrosis factor polypeptide, useful for treating
PT diseases, e.g. endotoxemic shock, inflammation, rheumatoid arthritis, or
PT inflammatory bowel disease.
XX
PS Disclosure; SEQ ID NO 9; 117pp; English.
XX
XX The invention relates to a novel isolated Drosophila tumor necrosis
CC factor (DmTNF) polypeptide. An isolated Drosophila tumor necrosis factor
CC (DmTNF) polypeptide comprises a polypeptide sequence selected from: (i)
CC an isolated polypeptide comprising amino acids 1-409, 2-409, 53-409, 62-
CC 409 of a sequence comprising 409 amino acids (AED55183), or (ii) an
CC isolated polypeptide comprising 332 contiguous amino acids of (AED55183),
CC where the polypeptide induces apoptosis in a cell in which the
CC polypeptide is recombinantly expressed, where the isolated polypeptide is
CC produced by a method, which comprises culturing an isolated recombinant
CC host cell comprising a vector comprising the coding region encoding the
CC polypeptide under conditions such that the polypeptide is expressed and
CC recovering the polypeptide. A polypeptide of the invention has
CC antibacterial, immunosuppressive, antiinflammatory, cytostatic,
CC anorectic, antidiabetic, antirheumatic, antiarthritic, and
CC gastrointestinal-gen. activity. The polypeptide, nucleic acids,
CC composition, and methods of the invention are useful for treating
CC diseases, e.g. endotoxemic shock, inflammation, hemorrhagic necrosis of
CC tumors, cytotoxicity, obesity-linked insulin resistance, rheumatoid
CC arthritis, or inflammatory bowel disease. The present sequence represents
CC murine ectodysplasmin A protein.
XX
SQ Sequence 391 AA;
Query Match 93.7%; Score 1965; DB 9; Length 391;
Best Local Similarity 94.6%; Pred. No. 1.4e-144;
Matches 370; Conservative 1; Mismatches 20; Indels 0; Gaps 0;
QY 1 MGYPEVERRELLPAAAPRRGSGCGCGAPARAGEGNSCLFLGPFGLSLALHLTLCC 60
DB 1 MGYPEVERREPLPAAAPRRGSGCGCGAPARAGEGNSCLFLGPFGLSLALHLTLCC 60
QY 61 YLELSELRRERGAESRLGSGTGTGTSGLSSGLDPSITSHLGQSPKQOPLERGE 120
DB 61 YLELSELRRERGTESRLGGPGAGTGTSLSSPGSLDPVGPITRHLGQSFQOQPLERGE 120
QY 121 AALHSDSQDGHQWALLNFFPFDEKPYSEESRRVRNRKSKNEGADGPVNKKKKKAG 180
DB 121 DPLPPDSQDRHQWALLNFFPFDEKAYSEESRRVRNRKSKSGEGADGPVNKKKKKAG 180
QY 181 PFGNPGPPGPPGPPGQPGIPGIPGIPGTTVMGPPGPPGPPGPPGPPGPPGPPGPPG 240
DB 181 PFGNPGPPGPPGPPGQPGIPGIPGIPGTTVMGPPGPPGPPGPPGPPGPPGPPGPPG 240
QY 241 ACTRENQRAVHLQGGSAIOVKNDLSGVNDWSRITWNPVKVFKLHPSRGEVLVDGT 300
DB 241 TGTRENQRAVHLQGGSAIOVKNDLSGVNDWSRITWNPVKVFKLHPSRGEVLVDGT 300
QY 301 YFIYSQVEVYVYINFTDFASEYVVDKPFELQCTRSJETGKNTYNTCYTAGVCLLKARQKI 360
DB 301 YFIYSQVEVYVYINFTDFASEYVVDKPFELQCTRSJETGKNTYNTCYTAGVCLLKARQKI 360
QY 361 AVKMVHADISINMSKHTTFFGAIRLGEAPAS 391
DB 361 AVKMVHADISINMSKHTTFFGAIRLGEAPAS 391

RESULT 10

ABB09083
ID ABB09083 standard; protein; 377 AA.
XX
AC ABB09083;
XX
DT 26-JUN-2002 (first entry)
XX
DE Mouse Tabby (Ta) protein SEQ ID NO:4.
XX
KW Human; mouse; EDAL-II; X-linked hypohidrotic ectodermal dysplasia; XLHED;
KW DL receptor; DL gene; downless gene; ectodermal dysplasia;
KW hair follicle growth; tooth growth; endocrine sweat gland development;
KW mammary epithelial tissue growth; breast epithelium.
XX
OS Mus musculus.
XX
PN US6355782-B1.
XX
PD 12-MAR-2002.
XX
PF 29-JUN-1999; 99US-00342681.
XX
PR 09-JUL-1998; 98US-0092279P.
PR 15-DEC-1998; 98US-0112366P.
XX
PA (BAYU) BAYLOR COLLEGE MEDICINE.
PA (UYOR-) UNIV OREGON HEALTH SCI.
XX
PI Zonana J, Ferguson BM, Headon D, Overbeek P;
DR WPI; 2002-360478/39.
DR N-PSDB; ABL51010.
XX
PT Novel purified hypohidrotic ectodermal dysplasias protein, useful for
PT promoting the development of hair follicles and tooth growth.
XX
PS Example 1; Col 71-74; 86pp; English.
XX
CC The present invention describes a purified hypohidrotic ectodermal
CC dysplasias protein (I) that promotes the development of hair follicles.
CC (I) is a promoter of hair follicle growth and tooth growth. (I) is useful
CC for the development of hair follicles. (I) is also useful for commercial
CC and clinical diagnostic testing for ectodermal dysplasia (ED). (I) is
CC also useful as a stimulant for hair and tooth growth, for stimulating
CC endocrine sweat gland development in individuals from whom the normal
CC sweating mechanism is compromised by disease or surgery, for stimulating
CC the growth of mammary epithelial tissue, either for reconstructing or
CC cosmetic purposes, and for promoting or maintaining differentiation of
CC breast epithelium. (I) is also useful for functional analysis, antibody
CC production and patient therapy. ABL51009 to ABL51122 and ABB09082 to
CC ABB09090 represent sequences used in the exemplification of the present
CC invention
XX
SQ Sequence 377 AA;
Query Match 89.3%; Score 1873; DB 5; Length 377;
Best Local Similarity 90.8%; Pred. No. 2e-137;
Matches 355; Conservative 2; Mismatches 20; Indels 14; Gaps 1;
QY 1 MGYPEVERRELLPAAAPRRGSGCGCGAPARAGEGNSCLFLGPFGLSLALHLTLCC 60
DB 1 MGYPEVERREPLPAAAPRRGSGCGCGAPARAGEGNSCLFLGPFGLSLALHLTLCC 60
QY 61 YLELSELRRERGAESRLGSGTGTGTSGLSSGLDPSITSHLGQSPKQOPLERGE 120
DB 61 YLELSELRRERGTESRLGGPGAGTGTSLSSPGSLDPVGPITRHLGQSFQOQPLERGE 120
QY 121 AALHSDSQDGHQWALLNFFPFDEKPYSEESRRVRNRKSKNEGADGPVNKKKKKAG 180
DB 121 DPLPPESQDRHQWALLNFFPFDEKAYSEESRRVRNRKSKSGEGADGPVNKKKKKAG 180
QY 181 PFGNPGPPGPPGPPGQPGIPGIPGIPGTTVMGPPGPPGPPGPPGPPGPPGPPGPPG 240


```
RESULT 14
AAW44764
XX AAW44764 standard; protein; 135 AA.
AC AAW44764;
XX
XX 01-JUN-1998 (first entry)
DT
DE Human anhidrotic ectodermal dysplasia protein.
XX
XX Human; anhidrotic ectodermal dysplasia; X chromosome; genetic linkage;
KW translocation; CpG island; foetal development; hair; sweat gland; tooth.
XX
XX Homo sapiens.
OS
XX
XX US5700926-A.
PN
XX
XX 23-DEC-1997.
PD
XX
XX 22-JUL-1996; 96US-00684672.
PF
XX
XX 27-APR-1993; 93US-00052997.
PR
XX
XX (UNIW ) UNIV WASHINGTON.
PA
XX
XX De La Chapelle A, Srivastava AK, Kere J, Schlessinger D;
PI
XX WPI; 1998-062436/06.
DR
XX N-PSDB; AAV05851.
DR
XX
XX Human anhidrotic ectodermal dysplasia gene - useful for research into
PT hair growth.
PT
XX
XX Disclosure; Fig 7; 37pp; English.
PS
XX
XX This is the amino acid sequence of the protein encoded by the human
CC anhidrotic ectodermal dysplasia (EDA) gene. The gene has been mapped to
CC the region Xq12-q13 by genetic linkage analysis and has been shown to
CC contain a 200 kb intron inserted in the 3' end of the coding sequence.
CC Deficiencies in the gene are observed by translocations with a breakpoint
CC in the transcribed CpG island 3 at the Xq12-q13 locus. The EDA gene can
CC be used to study the dynamics of EDA gene expression during foetal
CC development, and processes affecting normal hair growth in adults. The
CC EDA gene can also be used to study hair, sweat gland and tooth formation
CC and growth, and ectodermal dysplasias
XX
XX Sequence 135 AA;
SQ
Query Match 33.4%; Score 700; DB 2; Length 135;
Best Local Similarity 100.0%; Pred. No. 1.6e-46;
Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGYPEVERRELLPAAAPRRGSGCGCGGAPAPAGEGNSCLLFLGFFGLSLALHLTLCC 60
Db 1 MGYPEVERRELLPAAAPRRGSGCGCGGAPAPAGEGNSCLLFLGFFGLSLALHLTLCC 60
QY 61 YLELSELRRGAESRLGSGTPTGTSCTLSLGGLDPSITSHLGQSPKQOPLPGGE 120
Db 61 YLELSELRRGAESRLGSGTPTGTSCTLSLGGLDPSITSHLGQSPKQOPLPGGE 120
QY 121 AALHSDSDQGHQ 132
Db 121 AALHSDSDQGHQ 132
RESULT 15
AAR95115
ID AAR95115 standard; protein; 357 AA.
XX
XX AAR95115;
AC
XX
XX 25-MAR-2003 (revised)
DT
XX 03-FEB-1997 (first entry)
DT
```

```
XX
DE Intervening sequence contg. collagen like protein (CLP)-CB.
XX
KW Polymer; repeat unit; natural collagen; intervening oligopeptide; fibre;
KW film; membrane; emulsion; coating; collagen like protein;
KW specific binding material; catalyst; purification agent; composite;
KW laminate; adhesive; cell growth surface; affinity column;
KW biological material support; wound dressing; in vivo prothesis.
XX
OS Synthetic.
XX
XX US5514581-A.
PN
XX
XX 07-MAY-1996.
PD
XX
XX 06-NOV-1990; 90US-00609716.
PF
XX
XX 04-NOV-1986; 86US-00927258.
PR
XX 29-OCT-1987; 87US-00114618.
PR
XX 09-NOV-1988; 88US-00269429.
PR
XX 07-NOV-1989; 89WO-US005016.
XX
XX (PROT-) PROTEIN POLYMER TECHNOLOGIES INC.
PA
XX
XX Cappello J, Ferrari FA;
PI
XX WPI; 1996-238772/24.
DR
XX
XX DNA encoding protein contg. repeated fibroin derived segments - linked by
PT oligopeptide with cell adhesion properties useful, e.g. in wound
PT dressings.
PT
XX
XX Example 3; Col 101-104; 71pp; English.
PS
XX
XX Novel DNA sequence encodes a polymer comprising segments of repeating
CC units of 3-9 amino acids from natural collagen, i.e. the present
CC intervening sequence contg. collagen like protein (CLP)-CB, able to
CC assemble into aligned structures formable into articles. The polymer
CC comprises at least 2 segments joined by an unaligned intervening
CC oligopeptide, other than the repeating unit. The polymer can be used to
CC make fibres, films, membranes, emulsions, coatings, etc., useful as, e.g.
CC specific binding materials, catalysts, purificn. agents, composites,
CC laminates, adhesives, cell growth surfaces, affinity columns and supports
CC for biological materials. Typical applications include wound dressings,
CC and in vivo protheses. The polymer produces articles with good mechanical
CC properties, and the intervening oligopeptide can provide a ligand for
CC binding a mol., antibody, etc., or a chemically reactive site for
CC coupling to proteins, etc. (Updated on 25-MAR-2003 to correct PF field.)
XX
XX Sequence 357 AA;
SQ
Query Match 14.0%; Score 294; DB 2; Length 357;
Best Local Similarity 32.4%; Pred. No. 2.1e-14;
Matches 69; Conservative 9; Mismatches 71; Indels 64; Gaps 3;
QY 71 ERGASRLGGSGTPTGTSCTLSLGGLDPS--ITSHLGQSPKQOPLPFGAALHSDSQ 128
Db 71 DRGDAGPKGADSGPAGPVGSPGAPGPPGPGPPGPPGPPGPPGPPGPPGPPGPPG 130
QY 129 DGHQWALLNFFPFDEKPYSEESRVRNRKRSKNEGADGPVNKKKKGKAGPPGPN 188
Db 131 D-----RGDAGPKGADSGPAGPVGSPGAPGPPGPPGPP 165
QY 189 GPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP 212
Db 166 GPPGAPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP 225
QY 213 -VMGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP 244
Db 226 GPPGAPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP 258
Search completed: June 2, 2006, 17:24:59
```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 2, 2006, 17:25:22 ; Search time 41 Seconds
(without alignments)
917.579 Million cell updates/sec

Title: US-09-729-658C-2
Perfect score: 2098
Sequence: 1 MGYPEVERRELLPAAAPRR.....NMSKHTTFFGAIRLGEAPAS 391

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 80:*

1: Pirl:.*
2: Pirl2:.*
3: Pirl3:.*
4: Pirl4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	280.5	13.4	1838	1 CGHU1V	collagen alpha 1(V)
2	277	13.2	1464	2 S59856	collagen alpha 1(I)
3	275.5	13.1	1433	2 A46053	bullous pemphigoid
4	275	13.1	1049	1 CGB078	collagen alpha 1(I)
5	273.5	13.0	1843	2 S18803	collagen alpha 1(V)
6	273	13.0	1532	2 A61262	collagen alpha 1(X)
7	268.5	12.8	742	2 JCY595	scavenger receptor
8	267.5	12.8	316	2 T20497	hypothetical prote
9	267	12.7	886	2 I50694	collagen alpha 1(I)
10	267	12.7	1315	2 A56101	collagen alpha 1(X)
11	267	12.7	1774	2 B56101	collagen alpha 1(X)
12	266	12.7	779	1 CGB01S	collagen alpha 1(I)
13	265.5	12.7	1464	1 CGHU1S	collagen alpha 1(I)
14	265	12.6	1763	2 S16366	collagen alpha 2(I)
15	264	12.6	1690	1 CGBU1B	collagen alpha 4(I)
16	263.5	12.6	1146	2 A38587	collagen, cornea-s
17	263.5	12.6	1419	2 A41182	collagen alpha 1(I)
18	263.5	12.6	1487	2 B41182	collagen alpha 1(I)
19	263	12.5	1366	1 CGHU2S	collagen alpha 2(I)
20	260.5	12.4	675	2 S20819	collagen alpha 3(I)
21	259.5	12.4	302	2 T15936	hypothetical prote
22	259.5	12.4	674	2 S23297	collagen alpha 1(X)
23	258.5	12.3	1373	1 A43291	collagen alpha 2(I)
24	258	12.3	674	2 S13301	collagen alpha 1(X)
25	258	12.3	680	1 CGHU1D	collagen alpha 1(X)
26	258	12.3	1142	2 JX0369	collagen alpha 1(X)
27	257.5	12.3	312	2 T25048	hypothetical prote
28	257.5	12.3	1453	2 S21626	collagen alpha 1(I)
29	256.5	12.2	688	2 A53330	collagen alpha 2(I)

RESULT 1

CGHU1V

collagen alpha 1(V) chain precursor - human
N;Alternate names: procollagen alpha 1(V) chain
C;Species: Homo sapiens (man)

C;Date: 22-Nov-1993 #sequence revision 03-Oct-1995 #text change 09-Jul-2004
C;Accession: S18802; S16024; A61142; S11303; S03978; S43642; S58665
R;GreenSpan, D.S.; Cheng, W.; Hoffman, G.G.
J. Biol. Chem. 266, 24727-24733, 1991

A;Title: The pro-alpha1(V) collagen chain. Complete primary structure, distribution of e
A;Reference number: S18802; MUID:92105142; PMID:1722213
A;Accession: S18802
A;Molecule type: mRNA

A;Residues: 1-1838 <GRE>

A;Cross-references: UNIPARC:Q15094; UNIPARC:UPI0000065982; GB:M76729; NID:q189519; PIDN:R;Takahara, K.; Sato, Y.; Okazawa, K.; Okamoto, N.; Noda, A.; Yaoi, Y.; Kato, I.
J. Biol. Chem. 266, 13124-13129, 1991
A;Title: Complete primary structure of human collagen alpha-1(V) chain.
A;Reference number: S16024; MUID:91302336; PMID:2071595
A;Accession: S16024

A;Molecule type: mRNA

A;Residues: 1-81, 'QL', '84-389, 'A', '391-676, 'K', '678-1294, 'PS', '1297, 'RS', '1300-1553, 'R', '1555-1
A;Cross-references: UNIPARC:UPI0000126D1D; GB:D90279; NID:G219509; PIDN:BAAL4323.1; PIDN:R;Yaoi, Y.; Hashimoto, K.; Takahara, K.; Kato, I.
Exp. Cell Res. 194, 180-185, 1991
A;Title: Insulin binds to type V collagen with retention of mitogenic activity.
A;Reference number: A61142; MUID:91224163; PMID:1709100
A;Accession: A61142

A;Molecule type: protein

A;Residues: 823-824, 'X', '826-842 <YAO>

A;Cross-references: UNIPARC:UPI0000173BB7
A;Note: the residue designated 'X' is probably glycosylated hydroxylysine; this cyanogen
R;Yaol, Y.; Hashimoto, K.; Koitabashi, H.; Takahara, K.; Ito, M.; Kato, I.
Biochim. Biophys. Acta 1035, 139-145, 1990
A;Title: Primary structure of the heparin-binding site of type V collagen.
A;Reference number: S11303; MUID:90366601; PMID:2203476
A;Accession: S11303

A;Molecule type: protein

A;Residues: 823-824, 'X', '826-848, 'I', '850-851, 'P', '853, 'PR', '856-893, 'D', '895-932, 'X', '934-950

A;Cross-references: UNIPARC:UPI0000173BB8

A;Note: the residues designated 'X' are probably glycosylated hydroxylysine; this sequen
R;Seyer, J.M.; Kang, A.H.
Arch. Biochem. Biophys. 271, 120-129, 1989

A;Title: Covalent structure of collagen: amino acid sequence of three cyanogen bromide-d
A;Reference number: S03978; MUID:89227189; PMID:2496661
A;Accession: S03978

A;Molecule type: protein

A;Residues: 621-640, 'G', '642-649, 'L', '651-662, 'E', '664-667, 'Q', '669-676, 'Q', '678-683, 'P', '685-6
<SEY>

A;Cross-references: UNIPARC:UPI0000173BB9

A;Note: there are a number of inconsistencies between the sequences in figures 6 and 7; t

R;Moradi-Ameli, M.; Rouseau, J.C.; Kleman, J.P.; Champlaud, M.F.; Boutillon, M.M.; B&eur. J. Biochem. 221, 987-995, 1994

A:Title: Diversity in the processing events at the N-terminus of type-V collagen.

A:Reference number: S43642; MUID:94237164; PMID:8181482

A:Accession: S43642

A:Molecule type: protein

A:Residues: 565-576/756-758, 'X', 760-763, 'X', 765-772, 1012-1029, 1219-1232, 1465-1474, 'X', 1474-1482, 1565-1576/1577-1578, 'X', 1580-1583, 'X', 1586-1589, 'X', 1592-1595, 'X', 1598-1601, 'X', 1604-1607, 'X', 1610-1613, 'X', 1616-1619, 'X', 1622-1625, 'X', 1628-1631, 'X', 1634-1637, 'X', 1640-1643, 'X', 1646-1649, 'X', 1652-1655, 'X', 1658-1661, 'X', 1664-1667, 'X', 1670-1673, 'X', 1676-1679, 'X', 1682-1685, 'X', 1688-1691, 'X', 1694-1697, 'X', 1700-1703, 'X', 1706-1709, 'X', 1712-1715, 'X', 1718-1721, 'X', 1724-1727, 'X', 1730-1733, 'X', 1736-1739, 'X', 1742-1745, 'X', 1748-1751, 'X', 1754-1757, 'X', 1760-1763, 'X', 1766-1769, 'X', 1772-1775, 'X', 1778-1781, 'X', 1784-1787, 'X', 1790-1793, 'X', 1796-1799, 'X', 1802-1805, 'X', 1808-1811, 'X', 1814-1817, 'X', 1820-1823, 'X', 1826-1829, 'X', 1832-1835, 'X', 1838-1841, 'X', 1844-1847, 'X', 1850-1853, 'X', 1856-1859, 'X', 1862-1865, 'X', 1868-1871, 'X', 1874-1877, 'X', 1880-1883, 'X', 1886-1889, 'X', 1892-1895, 'X', 1898-1901, 'X', 1904-1907, 'X', 1910-1913, 'X', 1916-1919, 'X', 1922-1925, 'X', 1928-1931, 'X', 1934-1937, 'X', 1940-1943, 'X', 1946-1949, 'X', 1952-1955, 'X', 1958-1961, 'X', 1964-1967, 'X', 1970-1973, 'X', 1976-1979, 'X', 1982-1985, 'X', 1988-1991, 'X', 1994-1997, 'X', 1998-2000

A:Cross-references: UNIPARC:UPI0000173BBA; UNIPARC:UPI0000173BBB; UNIPARC:UPI0000173BBC; UNIPARC:UPI0000173BBD; UNIPARC:UPI0000173BBE; UNIPARC:UPI0000173BBF; UNIPARC:UPI0000173BBG; UNIPARC:UPI0000173BBH; UNIPARC:UPI0000173BBI; UNIPARC:UPI0000173BBJ; UNIPARC:UPI0000173BBK; UNIPARC:UPI0000173BBL; UNIPARC:UPI0000173BBM; UNIPARC:UPI0000173BBN; UNIPARC:UPI0000173BBO; UNIPARC:UPI0000173BBP; UNIPARC:UPI0000173BBQ; UNIPARC:UPI0000173BBR; UNIPARC:UPI0000173BBT; UNIPARC:UPI0000173BBU; UNIPARC:UPI0000173BBV; UNIPARC:UPI0000173BBW; UNIPARC:UPI0000173BBX; UNIPARC:UPI0000173BBY; UNIPARC:UPI0000173BBZ; UNIPARC:UPI0000173CBA; UNIPARC:UPI0000173CBF; UNIPARC:UPI0000173CBG; UNIPARC:UPI0000173CBH; UNIPARC:UPI0000173CBI; UNIPARC:UPI0000173CBJ; UNIPARC:UPI0000173CBK; UNIPARC:UPI0000173CBL; UNIPARC:UPI0000173CBM; UNIPARC:UPI0000173CBN; UNIPARC:UPI0000173CBP; UNIPARC:UPI0000173CBQ; UNIPARC:UPI0000173CBR; UNIPARC:UPI0000173CBS; UNIPARC:UPI0000173CBT; UNIPARC:UPI0000173CBU; UNIPARC:UPI0000173CBV; UNIPARC:UPI0000173CBW; UNIPARC:UPI0000173CBX; UNIPARC:UPI0000173CBY; UNIPARC:UPI0000173CBZ; UNIPARC:UPI0000173CCA; UNIPARC:UPI0000173CCF; UNIPARC:UPI0000173CCG; UNIPARC:UPI0000173CCH; UNIPARC:UPI0000173CCI; UNIPARC:UPI0000173CCJ; UNIPARC:UPI0000173CCK; UNIPARC:UPI0000173CCL; UNIPARC:UPI0000173CCM; UNIPARC:UPI0000173CCN; UNIPARC:UPI0000173CCO; UNIPARC:UPI0000173CCP; UNIPARC:UPI0000173CCQ; UNIPARC:UPI0000173CCR; UNIPARC:UPI0000173CCS; UNIPARC:UPI0000173CCT; UNIPARC:UPI0000173CCU; UNIPARC:UPI0000173CCV; UNIPARC:UPI0000173CCW; UNIPARC:UPI0000173CCX; UNIPARC:UPI0000173CCY; UNIPARC:UPI0000173CCZ; UNIPARC:UPI0000173CDA; UNIPARC:UPI0000173CDF; UNIPARC:UPI0000173CDG; UNIPARC:UPI0000173CDH; UNIPARC:UPI0000173CDI; UNIPARC:UPI0000173CDJ; UNIPARC:UPI0000173CDK; UNIPARC:UPI0000173CDL; UNIPARC:UPI0000173CDM; UNIPARC:UPI0000173CDN; UNIPARC:UPI0000173CDO; UNIPARC:UPI0000173CDP; UNIPARC:UPI0000173CDQ; UNIPARC:UPI0000173CDR; UNIPARC:UPI0000173CDS; UNIPARC:UPI0000173CDT; UNIPARC:UPI0000173CDU; UNIPARC:UPI0000173CDV; UNIPARC:UPI0000173CDW; UNIPARC:UPI0000173CDX; UNIPARC:UPI0000173CDY; UNIPARC:UPI0000173CDZ; UNIPARC:UPI0000173CEA; UNIPARC:UPI0000173CEF; UNIPARC:UPI0000173CEG; UNIPARC:UPI0000173CEH; UNIPARC:UPI0000173CEI; UNIPARC:UPI0000173CEJ; UNIPARC:UPI0000173CEK; UNIPARC:UPI0000173CEL; UNIPARC:UPI0000173CEM; UNIPARC:UPI0000173CEN; UNIPARC:UPI0000173CEO; UNIPARC:UPI0000173CEP; UNIPARC:UPI0000173CEQ; UNIPARC:UPI0000173CED; UNIPARC:UPI0000173CEE; UNIPARC:UPI0000173CEF; UNIPARC:UPI0000173CEG; UNIPARC:UPI0000173CEH; UNIPARC:UPI0000173CEI; UNIPARC:UPI0000173CEJ; UNIPARC:UPI0000173CEK; UNIPARC:UPI0000173CEL; UNIPARC:UPI0000173CEM; UNIPARC:UPI0000173CEN; UNIPARC:UPI0000173CEO; UNIPARC:UPI0000173CEP; UNIPARC:UPI0000173CEQ; UNIPARC:UPI0000173CED; UNIPARC:UPI0000173CEE; UNIPARC:UPI0000173CEF; UNIPARC:UPI0000173CEG; UNIPARC:UPI0000173CEH; UNIPARC:UPI0000173CEI; UNIPARC:UPI0000173CEJ; UNIPARC:UPI0000173CEK; UNIPARC:UPI0000173CEL; UNIPARC:UPI0000173CEM; UNIPARC:UPI0000173CEN; UNIPARC:UPI0000173CEO; UNIPARC:UPI0000173CEP; UNIPARC:UPI0000173CEQ; UNIPARC:UPI0000173CED; UNIPARC:UPI0000173CEE; UNIPARC:UPI0000173CEF; UNIPARC:UPI0000173CEG; UNIPARC:UPI0000173CEH; UNIPARC:UPI0000173CEI; UNIPARC:UPI0000173CEJ; UNIPARC:UPI0000173CEK; UNIPARC:UPI0000173CEL; UNIPARC:UPI0000173CEM; UNIPARC:UPI0000173CEN; UNIPARC:UPI0000173CEO; UNIPARC:UPI0000173CEP; UNIPARC:UPI0000173CEQ; UNIPARC:UPI0000173CED; UNIPARC:UPI0000173CEE; UNIPARC:UPI0000173CEF; UNIPARC:UPI0000173CEG; UNIPARC:UPI0000173CEH; UNIPARC:UPI0000173CEI; UNIPARC:UPI0000173CEJ; UNIPARC:UPI0000173CEK; UNIPARC:UPI0000173CEL; UNIPARC:UPI0000173CEM; UNIPARC:UPI0000173CEN; UNIPARC:UPI0000173CEO; UNIPARC:UPI0000173CEP; UNIPARC:UPI0000173CEQ; UNIPARC:UPI0000173CED; UNIPARC:UPI0000173CEE; UNIPARC:UPI0000173CEF; UNIPARC:UPI0000173CEG; UNIPARC:UPI0000173CEH; UNIPARC:UPI0000173CEI; UNIPARC:UPI0000173CEJ; UNIPARC:UPI0000173CEK; UNIPARC:UPI0000173CEL; UNIPARC:UPI0000173CEM; UNIPARC:UPI0000173CEN; UNIPARC:UPI0000173CEO; UNIPARC:UPI0000173CEP; UNIPARC:UPI0000173CEQ; UNIPARC:UPI0000173CED; UNIPARC:UPI0000173CEE; UNIPARC:UPI0000173CEF; UNIPARC:UPI0000173CEG; UNIPARC:UPI0000173CEH; UNIPARC:UPI0000173CEI; UNIPARC:UPI0000173CEJ; UNIPARC:UPI0000173CEK; UNIPARC:UPI0000173CEL; UNIPARC:UPI0000173CEM; UNIPARC:UPI0000173CEN; UNIPARC:UPI0000173CEO; UNIPARC:UPI0000173CEP; UNIPARC:UPI0000173CEQ; UNIPARC:UPI0000173CED; UNIPARC:UPI0000173CEE; UNIPARC:UPI0000173CEF; UNIPARC:UPI0000

Db 857 PLGPPGKGLGVDPGLPGVPRGQPKGS-----IGFPGFP----- 891
 Qy 72 RGAESRLGGSGTPTGTSGLSSLGGLDPDSFITSHLGGQSPKQQLPEFGAALHSDSQDGH 131
 Db 892 -GANGKGRGTP-----GKPGRGQGGTG----- 916
 Qy 132 QMALLNFFPDDEKPYSEESRRVRNRKSKSNEGADGVKNKKKKKAGP-----GPNGPP 188
 Db 917 -----PRGERGPRGITGPKGPKGNSSGDCP-----AGPPGERGPNQP 954
 Qy 189 GPPPPGPPGQGGTPIGIGIQ-----TTVMGPPGPPGP----- 222
 Db 955 GPTGFGGKGGPPGPKDGLPHGQRGGTGFGKTPGPPGPPGVGPGTGTGTGPMOER 1014
 Qy 223 --PGQPPGPGQGSAAADKAGTRENQPAVHLOQ 256
 Db 1015 GHPGPPGPGGGLPLAGKEGTK-GDQCPAGLPGK 1049

 RESULT 2
 S59856
 collagen alpha 1(III) chain precursor - mouse
 C;Species: Mus musculus (house mouse)
 C;Date: 10-Apr-1996 #sequence revision 19-Apr-1996 #text_change 09-Jul-2004
 C;Accession: S59856; S62120; S16373
 R;Toman, P.D.; de Crombrughe, B.
 Gene 147, 161-168, 1994
 A;Title: The mouse type-III procollagen-encoding gene: genomic cloning and complete DNA
 A;Reference number: S59856; MUID:95011609; PMID:7926795
 A;Accession: S59856
 A;Molecule type: DNA
 A;Residues: 1-1464 <TOM>
 A;Cross-references: UNIPROT:P08121; UNIPARC:UPI0000177386; EMBL:X52046
 R;Toman, D.
 submitted to the EMBL Data Library, November 1994
 A;Reference number: S62120
 A;Accession: S62120
 A;Molecule type: DNA
 A;Residues: 1-866, 'G', 868-1464 <TOM>
 A;Cross-references: UNIPARC:UPI0000295D6; EMBL:X52046; NID:g575321; PIDN:CAA36279.1; PID
 R;Metsaeranta, M.; Toman, D.; de Crombrughe, B.; Vuorio, E.
 Biochim. Biophys. Acta 1089, 241-243, 1991
 A;Title: Specific hybridization probes for mouse type I, II, III and IX collagen mRNAs.
 A;Reference number: S16176; MUID:91274355; PMID:2054384
 A;Accession: S16373
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1442-1464 <MET>
 A;Cross-references: UNIPARC:UPI000016CCAB; EMBL:X57983; NID:g50476; PIDN:CAA41048.1; PID
 C;Genetics:
 A;Introns: 29/1; 95/3; 112/3; 150/3; 175/3; 193/3; 211/3; 229/3; 247/3; 265/3; 283/3; 299/
 58/3; 673/3; 706/3; 742/3; 760/3; 778/3; 796/3; 814/3; 850/3; 868/3; 886/3; 940/3; 976/3;
 C;Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;
 C;Keywords: coiled coil; extracellular matrix
 F;1-24/Domain: signal sequence #status predicted <SIG>
 F;25-154/Domain: propeptide #status predicted <PRO>
 F;32-92/Domain: von Willebrand factor type C repeat homology <VWC>
 F;155-1464/Product: collagen alpha 1(III) chain #status predicted <MAT>
 F;1236-1464/Domain: fibrillar collagen carboxyl-terminal homology <FCC>

 Query Match 13.2%; Score 277; DB 2; Length 1464;
 Best Local Similarity 31.8%; Pred. No. 2.2e-10;
 Matches 76; Conservative 6; Mismatches 75; Indels 82; Gaps 8;

 Qy 13 PAAAPRRGSG-----CCCGAPAPRAGGNSCLLFLGFGLSLALHLLTLCCVLELRSEL 68
 Db 792 PRGGFGERGEPGPPGAPGPPGAGNGBPFA----- 822
 Qy 69 RRRGASRLGGSGTPTGTSGLSSLGGLDPDSF--ITSHLGGQSPKQQLPEFGAALHSD 126
 Db 823 KGERGAPGKEGGGPPGAPGTGSSGPPAGPPGQGVKGRGSGFP-----PGTAG----- 872
 Qy 127 SQDGHQMALNFFPFDEKPYSEESRRVRNRKSKSNEGADGVKNKKKKKAGPPGPN 186

[illegible]

R;Dewes, H.; Fietzek, P.P.; Kuehn, K.
Hoppe-Seyler's Z. Physiol. Chem. 360, 821-832, 1979
A;Title: The covalent structure of calf skin type III collagen. II. The amino acid sequence of the alpha 1(I) chain. PMID:488907
A;Reference number: A38001; MUID:80026027; PMID:488907
A;Accession: A38001
A;Molecule type: protein
A;Residues: 243-422 <DEWI>
A;Cross-references: UNIPARC:UPI0000173B88
R;Bentz, H.; Fietzek, P.P.; Kuehn, K.
Hoppe-Seyler's Z. Physiol. Chem. 360, 833-840, 1979
A;Title: The covalent structure of calf skin type III collagen. III. The amino acid sequence of the alpha 2(I) chain. PMID:488908
A;Reference number: A38002; MUID:80026028; PMID:488908
A;Accession: A38002
A;Molecule type: protein
A;Residues: 423-571 <BEN>
A;Cross-references: UNIPARC:UPI0000173B8C
R;Lang, H.; Glanville, R.W.; Fietzek, P.P.; Kuehn, K.
Hoppe-Seyler's Z. Physiol. Chem. 360, 841-850, 1979
A;Title: The covalent structure of calf skin type III collagen. IV. The amino acid sequence of the alpha 3(I) chain. PMID:488909
A;Reference number: A38003; MUID:80026029; PMID:488909
A;Accession: A38003
A;Molecule type: protein
A;Residues: 572-808 <LAN>
A;Cross-references: UNIPARC:UPI0000173B8D
R;Dewes, H.; Fietzek, P.P.; Kuehn, K.
Hoppe-Seyler's Z. Physiol. Chem. 360, 851-860, 1979
A;Title: The covalent structure of calf skin type III collagen. V. The amino acid sequence of the alpha 4(I) chain. PMID:488910
A;Reference number: A38004; MUID:80026030; PMID:488910
A;Accession: A38004
A;Molecule type: protein
A;Residues: 809-947 <DEW2>
A;Cross-references: UNIPARC:UPI0000173B8E
R;Allmann, H.; Fietzek, P.P.; Glanville, R.W.; Kuehn, K.
Hoppe-Seyler's Z. Physiol. Chem. 360, 861-868, 1979
A;Title: The covalent structure of calf skin type III collagen. VI. The amino acid sequence of the alpha 5(I) chain. PMID:488911
A;Reference number: A38005; MUID:80026031; PMID:488911
A;Accession: A38005
A;Molecule type: protein
A;Residues: 948-1049 <ALL>
A;Cross-references: UNIPARC:UPI0000173B8F
A;Experimental source: skin
R;Henkel, W.
Biochem. J. 318, 497-503, 1996
A;Title: Cross-link analysis of the C-telopeptide domain from type III collagen.
A;Reference number: S71946; MUID:96404897; PMID:8809038
A;Accession: S71946
A;Molecule type: protein
A;Residues: 81-106;1017-1029;1037-1049 <HEN>
A;Cross-references: UNIPARC:UPI0000173B90; UNIPARC:UPI0000173B91; UNIPARC:UPI0000173B92
C;Comment: Prolines at the third position of the tripeptide repeating unit (G-X-Y) are highly conserved. The type III collagen molecule is a trimer of identical chains, linked to each other by interchain disulfide bonds.
C;Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;
C;Keywords: coiled coil; extracellular matrix; glycoprotein; hydroxylysine; hydroxyproline
F;1-1049/Product: collagen alpha 1(I) chain #status experimental
F;1-14/Region: amino-terminal nonhelical telopeptide
F;15-1040/Region: helical
F;587-589/Region: cell attachment (R-G-D) motif
F;752-754/Region: cell attachment (R-G-D) motif
F;875-877/Region: cell attachment (R-G-D) motif
F;878-880/Region: cell attachment (R-G-D) motif
F;935-937/Region: cell attachment (R-G-D) motif
F;1041-1049/Region: carboxyl-terminal nonhelical telopeptide
F;95, 107, 119, 938, 950/Modified site: 5-hydroxylysine (Lys) #status experimental
F;107, 950/Modified site: allysine (Lys) #status predicted
F;107/Binding site: carbohydrate (Lys) (covalent) #status experimental
F;1040, 1041/Disulfide bonds: interchain #status predicted

Query Match 13.1%; Score 275; DB 1; Length 1049;
Best Local Similarity 31.3%; Pred. No. 2, 1e-10;
Matches 77; Conservative 5; Mismatches 68; Indels 96; Gaps 8;

Qy 13 PAAAPRRGSGG----CCCGAPAPAGSGNCLFLGFFGLSLAHLTLCCVLEURSEL 68

A;Molecule type: protein
A;Residues: 1-242 <PIE>
A;Cross-references: UNIPROT:P04258; UNIPARC:UPI0000173B8A

```
Db 637 PRGGPGERGEQQPPGAGPGAPQNGEPGA----- 667
QY 69 RRERGAESRLGSGTGTGTSLSLGLDPSDPSITSHLGQSPKQOPLRPGEAALHSDSQ 128
Db 668 KGERGAPGKEGEGPPGAGPAGCGSPAGPPG-----QGVKGERGSPGPGAG----- 717
QY 129 DGHOMALLNFFPFDKPYSEESRRVRNRKSKSNEGADGPVNKKKKKAGPGPNPGPP 188
Db 718 -----PB-----GGRGP-----PGPPGSGNGP 734
QY 189 GPP-----GPPGQPPGPIGPIGPIGT-VMGPPGPPGPPGPPGPPGLQGPSGAA 238
Db 735 GPPGSSGAPGKDGPPGPPGSGNAGPSPGSGKSDSGPPGGERGAPGQPPGAPGLGTA 794
QY 239 DRAGTR 244
Db 795 GLTGAR 800

RESULT 5
S18803
collagen alpha 1(V) chain - hamster
C:Species: Cricetineae gen. sp. (hamster)
C>Date: 19-Mar-1997 #sequence_revision 24-Jul-1997 #text_change 16-Dec-1998
C:Accession: S18803
R:GreenSPAN, D.S.; Cheng, W.; Hoffman, G.G.
J: Biol. Chem. 266, 24727-24733, 1991
A:Title: The pro-alpha1(V) collagen chain. Complete primary structure, distribution of e
A:Reference number: S18802; MUID:92105142; PMID:17221213
A:Accession: S18803
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1943 <RES>
A:Cross-references: UNIPARC:UPI00001773AC
C:Superfamily: collagen alpha 1(V) chain; fibrillar collagen carboxyl-terminal homology
F:1620-1842/Domain: fibrillar collagen carboxyl-terminal homology <FCC>

Query Match 13.0%; Score 273.5; DB 2; Length 1843;
Best Local Similarity 29.7%; Pred. No. 4.8e-10;
Matches 82; Conservative 9; Mismatches 70; Indels 115; Gaps 10;

QY 13 PAAAPRERSQGC-GCGGAPARAGENSCLLFLGFLGLSLALHLTLCCYLELRSELRR 71
Db 862 PLGPTCEKGLGVPLGPGVGRQPKGS-----IGPPGPF----- 896
QY 72 RGAESRLGSGTGTGTSLSLGLDPSDPSITSHLGQSPKQOPLRPGEAALHSDSQGH 131
Db 897 -GANGKEKGRGTP-----GKPGRQGRGPTG----- 921
QY 132 QMALLNFFPFDKPYSEESRRVRNRKSKSNEGADGPVNKKKKKAGKAGPP--GPNGPP 188
Db 922 -----PRGERGPRGITGKPGKNSGGDGP-----AGPPGERGNGPQ 959
QY 189 GPPGPPGQPPGPIGPIGIPG-----TTVMGPPGPPGP----- 222
Db 960 GPTGFGPKGPPGPKDGLGHPGQGTGTGQKTGPPGPPGPPGPPGPPGPPGPPGPP 1019
QY 223 --PGPPGPPGLQGPSGAADKAGTTRENPQAVVHLQQQ 256
Db 1020 GHPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP 1054

RESULT 6
A61262
collagen alpha 1(XVII) chain - human (fragment)
N:Alternate names: bullous pemphigoid 180K autoantigen BPAG2; bullous pemphigoid antigen
C:Species: Homo sapiens (man)
C>Date: 12-May-1994 #sequence_revision 12-Jul-1996 #text_change 09-Jul-2004
C:Accession: I56325; I55345; A61262
R:Giudice, G.J.; Emery, D.J.; Diaz, L.A.
J: Invest. Dermatol. 99, 243-250, 1992
A:Title: Cloning and primary structural analysis of the Bullous pemphigoid autoantigen,
A:Reference number: I56325; MUID:92381323; PMID:1324962
```

```
A:Accession: I56325
A:Status: preliminary; translated from GB/EMBL/DBBJ
A:Molecule type: mRNA
A:Residues: 1-1532 <RES>
A:Cross-references: UNIPROT:Q9UMDN9; UNIPARC:UPI000006F673; GB:M91669; NID:g179516; PIDN:I
R:Li, K.H.; Sawamura, D.; Giudice, G.J.; Diaz, L.A.; Mattel, M.G.; Chu, M.L.; Utto, J.
J. Biol. Chem. 266, 24064-24069, 1991
A:Title: Genomic organization of collagenous domains and chromosomal assignment of human
A:Reference number: I55345; MUID:92084712; PMID:1748679
A:Accession: I55345
A:Status: preliminary; translated from GB/EMBL/DBBJ
A:Molecule type: mRNA
A:Residues: 543-890, 'P' <RE2>
A:Cross-references: UNIPARC:UPI000006F72A; GB:M63730; NID:g179520; PIDN:AAA51839.1; PID:
R:Giudice, G.J.; Squigera, H.L.; Elias, P.M.; Diaz, L.A.
J. Clin. Invest. 87, 734-738, 1991
A:Title: Identification of two collagen domains within the bullous pemphigoid autoantigen
A:Reference number: A61262; MUID:91123476; PMID:1846881
A:Accession: A61262
A:Molecule type: mRNA
A:Residues: 543-890, 'P' <GIU>
A:Cross-references: UNIPARC:UPI000006F72A
C:Genetics:
A:Gene: GDB:COL17A1; BPAG2; BP180
A:Cross-references: GDB:131396; OMIM:113811
A:Map position: 10q24.3-10q24.3

Query Match 13.0%; Score 273; DB 2; Length 1532;
Best Local Similarity 31.2%; Pred. No. 4.3e-10;
Matches 85; Conservative 13; Mismatches 64; Indels 110; Gaps 13;

QY 20 RGSQGC---GCGGAPARAGENSCLLFLGFLGLSLALHLTLCCYLELRSELRRERGA 75
Db 789 RGEQLTGMPGIRGPPGSGDP-----GKPGLT-----GPG 819
QY 76 SRLGSGTGTGTSLSLGLDPSDPSITSH-----LGQSPKQOPLRPG 119
Db 820 GPGPLGPTGPRGPKGEPGA--PGKIVTSEGSSMLTVPGPPGPPGPPGPPGPPG 877
QY 120 EAALHSDSQDGHOMALLNFFPFDKPYSEESRRVRNRKSKSNEGADGPVNKKKKKKA 179
Db 878 PAGL-----PGHQEVL-----NLQGP-----P 894
QY 180 GPPGPPGPPGPP--PGPPGPPGPP--GIPGIPG-----TTVMGPPGPPGPPGPP 225
Db 895 GPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP 954
QY 226 QGPPGLQGPSGAADKAG--TRENQPAVVHLQ 255
Db 955 QGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP 986

RESULT 7
JC7595
scavenger receptor with C-type lectin type I - human
C:Species: Homo sapiens (man)
C>Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 09-Jul-2004
C:Accession: JC7595
R:Nakamura, K.; Funakoshi, H.; Miyamoto, K.; Tokunaga, F.; Nakamura, T.
Biochem. Biophys. Res. Commun. 280, 1028-1035, 2001
A:Title: Molecular cloning and functional characterization of a human scavenger receptor
A:Reference number: JC7595; MUID:21092718; PMID:11162630
A:Contents: Placenta
A:Accession: JC7595
A:Molecule type: mRNA
A:Residues: 1-742 <NAK>
A:Cross-references: UNIPROT:Q9BYH7; UNIPARC:UPI000006F955; DBJ:AB038518
C:Comment: This receptor, a member of the scavenger receptor family, belonging to the ty
important role in host defense. It forms a timer and plays a role in recognizing infecte
C:Genetics:
A:Gene: src1-1
A:Map position: 18p11.32
C:Keywords: coiled coil; glycoprotein; transmembrane protein
```


Eur. J. Biochem. 30, 169-183, 1972

A:Title: The covalent structure of collagen. The amino-acid sequence of the 112 residues
A:Reference number: A91201; MUID:73042276; PMID:4343808

A:Accession: A91201

A:Molecule type: protein

A:Residues: 563-675 <WEN>

A:Cross-references: UNIPARC:UPI0000173B55

A:Experimental source: skin

R:Fietzek, P.P.; Rexrodt, F.W.; Wendt, P.; Stark, M.; Kuehn, K.

Eur. J. Biochem. 30, 163-168, 1972

A:Title: The covalent structure of collagen. Amino acid sequence of peptide alpha1-CB6-C;
A:Reference number: A91200; MUID:73042275; PMID:4343807

A:Accession: A91200

A:Molecule type: protein

A:Residues: 676-758 <PI4>

A:Cross-references: UNIPARC:UPI0000173B56

A:Experimental source: skin

A:Note: Pro-726 is the only 3-hydroxyproline and the only hydroxylated proline in position 726.

R:Auterberg, J.; Fietzek, P.; Rexrodt, F.; Becker, U.; Stark, M.; Kuehn, K.

FEBS Lett. 21, 75-79, 1972

A:Title: The amino acid sequence of the carboxyterminal nonhelical cross link region of t
A:Reference number: A43048

A:Accession: A43048

A:Molecule type: protein

A:Residues: 759-779 <RA2>

A:Cross-references: UNIPARC:UPI0000173B57

A:Experimental source: skin

C:Comment: Lysines at positions 115, 124, 274, 346, 424, 496, 658, and 670 may be hydroxy
C:Comment: Prolines in the third position of the tripeptide repeating unit (G-X-Y) are hy
C:Comment: The order of the eight CNBr peptides in the alpha 1(I) chain of bovine skin c
9, 149, 268, and 217 residues.

C:Comment: The complete chain contains 1052 residues.

C:Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;
C:Keywords: coiled coil; extracellular matrix; glycoprotein; pyroglyutamic acid; trimer;
P:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 12.7%; Score 266; DB 1; Length 779;
Best Local Similarity 30.0%; Pred. NO. 6.1e-10;
Matches 78; Conservative 16; Mismatches 80; Indels 86; Gaps 9;

	QY	DB	QY	DB	QY	DB	QY	DB	QY	DB	
13	PARAPRSGSG-----CGCGAPARAGE-----GNSCLFLGFGLSLALHLLTLCYLLEL 64	339	PAGAPDKGEAGPSGPAGTRGAPDGRGGPPGPA-----GFAG----- 377	65	RSELRERGAESRLGGSGTPGTSTLSLGLLDPOSPITSHLGQSPKQPQLPEFGAALH 124	378	-----PPGADQGQAKGEPSDAGKGDAGPPGPAGP-----AGPPGPIGNVGAGPP- 423	125	SDSQDHQMALLNFFPFDEKPYSEESRRVRNRKSKSNEGADGVPNKKKGKAGPGPP 184	424	-----KGARGSAGPPGATGFPFGAAGRVPVGGP 450
185	NGPPGPPGPPG-----QGPPGIPGIPGTTVMGPPGPPGPPGPPGPPGLOGPSGAA 238	451	SGNAGPPGPPGPKAGKEGSKPRGETGAPRGCE--VGPPGPPGPKAGKAGADGAPGAP 508	239	DKACTR--ENQPAVVHLQQ 256	509	GTPGQGIAGQRGVVGLPGQ 528				

RESULT 13

CGHULS

collagen alpha 1(I) chain precursor - human

N:Alternate names: procollagen alpha 1(I) chain

C:Species: Homo sapiens (man)

C:Date: 12-Aug-1991 #sequence revision 04-Oct-1996 #text change 31-Dec-2004

C:Accession: I60114; S01143; A93335; I55254; A39943; I55237; A35233; S09400; B90567; S11:

5269; A29439; I53466; A02852; I37247

R:D'Alessio, M.; Bernard, M.; Pretorius, P.J.; de Wet, W.; Ramirez, F.; Pretorius, P.J.

Gene 67, 105-115, 1998

A:Title: Complete nucleotide sequence of the region encompassing the first twenty-five e;
A:Reference number: I60114; MUID:88329734; PMID:2843432

A;Reference number: A35336; MUID:90252792; PMID:2339700
A;Accession: A35336
A;Molecule type: mRNA
A;Residues: 710-720,'E',722-737,'E',739-745 <WAL>
A;Cross-references: UNIPARC:UPI0000173B4B
A;Note: the authors translated the codons CAG for 721 and CGT for 738 as Glu
R;Forlino, A.; Zolezzi, F.; Valli, M.; Pignatti, P.F.; Cetta, G.; Brunelli, P.C.; Motte-
Hum. Mol. Genet. 3, 2201-2206, 1994
A;Title: Severe (type III) osteogenesis imperfecta due to glycine substitutions in the α 1(I) chain of ty-
A;Reference number: I54365; MUID:95187161; PMID:7881420
A;Accession: I54365
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 746-766,'S',768-781 <FOR>
A;Cross-references: UNIPARC:UPI000016A6FA; GB:I47667; NID:g1009093; PIDN:AAB59576.1; PID:
R;Chesler, S.D.; Wallis, G.A.; Byers, P.H.
J. Biol. Chem. 268, 18218-18225, 1993
A;Title: Mutations in the carboxyl-terminal propeptide of the pro α 1(I) chain of ty-
A;Reference number: A47426; MUID:93352646; PMID:8349697
A;Accession: A47426
A;Molecule type: mRNA
A;Residues: 1179-1276,'H',1278-1336,1339-1387,'R',1389-1464 <CH>
A;Cross-references: UNIPARC:UPI000073A2A; GB:S64596; NID:9407589; PIDN:AAB27856.1; PID:
A;Note: sequence extracted from NCBI backbone (NCBIN:136444, NCBI:136445)
A;Note: does not represent an experimentally determined sequence but three different mut-
A;Accession: B47426
A;Molecule type: mRNA
A;Residues: 1179-1464 <CH4>
A;Cross-references: UNIPARC:UPI0000173B4C
A;Experimental source: normal dermal fibroblast culture
A;Accession: C47426
A;Molecule type: mRNA
A;Residues: 1179-1276,'H',1278-1276,'H',1278-1464 <CH5>
A;Cross-references: UNIPARC:UPI0000173B4F
A;Experimental source: fetal cell 86-237
A;Accession: D47426
A;Molecule type: mRNA
A;Residues: 1179-1336,1339-1464 <CH6>
A;Cross-references: UNIPARC:UPI0000173B4E
A;Experimental source: fetal cell 86-146
A;Accession: E47426
A;Molecule type: mRNA
A;Residues: 1179-1387,'R',1389-1464 <CH7>
A;Cross-references: UNIPARC:UPI0000173B4F
A;Experimental source: fetal cell 88-251
R;Cohn, D.H.; Apone, S.; Eyre, D.R.; Starman, B.J.; Andreassen, P.; Charbonneau, H.; Nic-
J. Biol. Chem. 263, 14605-14607, 1988

Query Match 12.7%; Score 265.5; DB 1; Length 1464;
Best Local Similarity 29.7%; Pred. No. 1.3e-09;
Matches 78; Conservative 16; Mismatches 80; Indels 89; Gaps 9;

QY 13 PAAAPRRSGQ-----CGCGGAPARAGE-----GNSCLLFLGFFGLSLHLTLCCY 61
DB 774 PACAPGDKESGPSGAGPTGARGAPDRGEPGPPGA-----GFAG----- 815
QY 62 LEILSELRRERGAESRLGSGTGPTGTLSSLGGLDPDPSITSLHGLQSPKQOPLPGBA 121
DB 816 -----PPGADGGQPGAKGEPGDAGAKDAGPPGAPG-----AGPPGPIGNVGAPGA- 861
QY 122 ALHSDSQDHQMALNFFPDPEPYEESSRRVRNKRKSNEGADGVPVKNKKKKAGP 181
DB 862 -----KGARGSAGPPGATGFFGAAGRVP 885
QY 182 PGNGPGPPGPPGP-----QPPGPTGPIGPIGTVMGPPGPPGPPGPPGLQGPS 235
DB 886 PGFSGNAGPPGPPGPKAGCKGPRGETGAPRGE--VGPPGPPGAGKESFGADGPA 943
QY 236 GAADKAGTR--ENQPAVVHLQQ 256
DB 944 GAPGTPGQGIAGQGVVGLPGQ 966

RESULT 14

S16366
collagen alpha 2(IV) chain precursor - pig roundworm
C/Species: Ascaris suum (pig roundworm)
C/Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 09-Jul-2004
C/Accession: S16366
R/Pettitt, J.; Kingston, I.B.
J. Biol. Chem. 266, 16149-16156, 1991
A/Title: The complete primary structure of a nematode alpha-2(IV) collagen and the partial
A/Reference number: S16366; MUID:91340768; PMID:1714907
A/Accession: S16366
A/Molecule type: mRNA
A/Residues: 1-1763 <JBI>
A/Cross-references: UNIPROT:P27393; UNIPARC:UPI0000126D40; GB:M67507; NID:G159648; PIDN:J
C/Genetics:
A/Supernotes: 229/3; 266/3; 305/3; 360/3; 424/1; 489/1; 548/1; 656/3; 790/1; 891/1; 963/1; 1
C/Superfamily: collagen alpha 1(IV) chain
C/Keywords: alternative splicing; basement membrane; cell binding; coiled coil; disulfide
F/1-26/Domain: signal sequence #status predicted <SIG>
F/27-1763/Product: collagen alpha 2(IV) chain #status predicted <MAT>
F/27-42/Domain: non-collagenous NHI #status predicted <NH1>
F/43-1529/Domain: collagenous #status predicted <COL>
F/197-199/Region: cell attachment (R-G-D) motif
F/1530-1763/Domain: carboxyl-terminal nonhelical, NCI #status predicted <NC1>
F/1530-1638/Domain: repeat NCI #status predicted <NC11>
F/1639-1763/Domain: repeat NCI #status predicted <NC12>
F/31,34,39,41,536,539/Disulfide bonds: interchain #status predicted
F/126/Binding site: carbohydrate (Asn) (covalent) #status predicted
F/1593-1599,1702-1709/Disulfide bonds: #status predicted

Query Match 12.68; Score 265; DB 2; Length 1763;
Best Local Similarity 34.3; Pred. No. 1.6e-09;
Matches 85; Conservative 15; Mismatches 82; Indels 66; Gaps 10;

QY	13	PAAAPRRGSGQ-----CCGCGAPARAGSGNSCLLFLGFFGLSLALHLTLCCYLELR	65
DB	256	PEGAPGKMGKEGDPGEAGRGFPPTGVAGQP-----GLPGMKGERGL	298
QY	66	SELRERKGAESRLGGSTGPTGTSGLTSLSGGLDDPSITSHLGQSPKQOPLPGEAALHS	125
DB	299	SGPAGPRGKEGRPLGPPGPKGB-----RGLD-----GLPGVPGLPGQKEAGF--	343
QY	126	DSODGHQMALNFFFPDEKPVSEESRRVRNKRKSKSNEGADGPVKNNKKKKAGPPGPN	185
DB	344	PGRDGAKGAR-----GPPPGPGGGGFS-----DGGPPGPGLGREGQPGPPGAD	387
QY	186	GPPGPPGPPGPPGPPGPGIPGIPGIPGTTVMGPPGP-----PGPPGPGPPGPGSPG	236
DB	388	GYGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG	445
QY	237	AADKAGTR 244	
DB	446	AKGEPGPR 453	

RESULT 15

CGHUIB
collagen alpha 4(IV) chain precursor - human
N/Alternate names: procollagen alpha 4(IV) chain
C/Species: Homo sapiens (man)
C/Date: 06-Feb-1995 #sequence_revision 03-Oct-1995 #text_change 09-Jul-2004
C/Accession: A55360; S38854; S28777
R/Leinonen, A.; Marlyama, M.; Mochizuki, T.; Tryggvason, K.; Reeders, S.T.
J. Biol. Chem. 269, 26172-26177, 1994
A/Title: Complete primary structure of the human type IV collagen alpha4(IV) chain. Comp
A/Reference number: A55360; MUID:95014445; PMID:7523402
A/Accession: A55360
A/Status: nucleic acid sequence not shown
A/Molecule type: mRNA
A/Residues: 1-1690 <LEI>
A/Cross-references: UNIPROT:P53420; UNIPARC:UPI0000000769; GB:X81053; NID:G574805; PIDN:G
R/Sugimoto, M.; Oohashi, T.; Yoshioka, H.; Matsuo, N.; Ninomiya, Y.

GenCore version 5.1.9

Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 2, 2006, 17:21:42 ; Search time 298 Seconds
(without alignments)
1213.695 Million cell updates/sec

Title: US-09-729-658C-2

Perfect score: 2098

Sequence: 1 MGYPEVERRELLPAAAPRR.....NMSKHTTFGATRLGEAPAS 391

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot 7.2.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2098	100.0	391	1	EDA_HUMAN	Q92838 homo sapien
2	1972	94.0	391	1	EDA_BOVIN	Q9beg5 bos taurus
3	1965	93.7	391	1	EDA_MOUSE	O54693 mus musculus
4	1965	93.7	391	2	Q3UV69_MOUSE	Q3uv69 mus musculus
5	1890.5	90.1	386	2	Q5BP77_CANFA	Q5bp77 canis faml
6	1441.5	68.7	356	2	O5EFZ5_CHICK	Q5efz5 gallus gall
7	955.5	45.5	367	2	Q5OD54_GASAC	Q5od54 gasteroste
8	887	42.3	178	2	Q5JS00_HUMAN	Q5j800 homo sapien
9	809	38.6	207	2	Q4SJR1_TETNG	Q4sjr1 tetraodon n
10	286	13.6	742	2	Q8VIF6_MOUSE	Q8vif6 mus musculus
11	284	13.5	998	2	Q8CFM4_MOUSE	Q8cfm4 mus musculus
12	282	13.4	742	2	Q8C979_MOUSE	Q8c979 mus musculus
13	282	13.4	742	2	Q8K4Q8_MOUSE	Q8k4q8 mus musculus
14	280.5	13.4	1792	2	Q59EE7_HUMAN	Q59ee7 homo sapien
15	280.5	13.4	1838	1	COSA1_HUMAN	P20908 homo sapien
16	280.5	13.4	1840	2	Q59IP3_PIG	Q59ip3 sus scrofa
17	280	13.3	284	2	Q25581_9BILA	Q25581 teladorsagi
18	278	13.3	742	2	Q4V985_RAT	Q4v985 ratus norv
19	277.5	13.2	1838	1	COSA1_MOUSE	Q88207 mus musculus
20	277.5	13.2	1840	1	COSA1_RAT	Q9j103 ratus norv
21	276	13.2	622	2	Q9BY85_HUMAN	Q9by85 homo sapien
22	276	13.2	1315	2	Q8QHL9_XENLA	Q8qhl9 xenopus lae
23	275.5	13.1	1470	1	COHAI_MOUSE	Q07563 mus musculus
24	275	13.1	1049	1	CO3A1_BOVIN	P04258 bos taurus
25	275	13.1	1431	1	COHAI_MESAU	Q9jfm4 mesocricetu
26	273.5	13.0	1840	1	COSA1_CRILO	Q60467 cricetus
27	273	13.0	742	2	Q5KU26_HUMAN	O5ku26 homo sapien
28	273	13.0	742	2	Q8WZA4_HUMAN	O8wza4 homo sapien
29	273	13.0	1497	1	COHAI_HUMAN	Q9umd9 homo sapien
30	272	13.0	284	2	Q25582_9BILA	Q25582 teladorsagi
31	272	13.0	1415	2	Q5UV36_HUMAN	Q5uv36 homo sapien

32	272	13.0	1802	2	Q17163_BRUMA	Q17163 brugia mala
33	271.5	12.9	1344	2	O93419_CHICK	O93419 gallus gall
34	271	12.9	367	2	O5CIA3_SCHJA	O5cia3 schistosoma
35	270	12.9	1222	2	Q8K173_MOUSE	Q8k173 mus musculus
36	270	12.9	1464	1	CO3A1_MOUSE	.P08121 mus musculus
37	270	12.9	1464	2	Q3TVI5_MOUSE	Q3tvi5 m osteoclas
38	270	12.9	1464	2	Q3UH72_MOUSE	Q3uh72 mus musculus
39	270	12.9	1464	2	Q7TT32_MOUSE	Q7tt32 mus musculus
40	270	12.9	1464	2	Q8BLW4_MOUSE	Q8blw4 mus musculus
41	270	12.9	1467	2	Q5DTG2_MOUSE	Q5dtg2 mus musculus
42	268.5	12.8	742	2	Q6P9F2_HUMAN	Q6p9f2 homo sapien
43	268.5	12.8	742	2	Q9BYH7_HUMAN	Q9byh7 homo sapien
44	268.5	12.8	1258	2	Q8AW11_BRARE	Q8aw11 brachydanio
45	268.5	12.8	1835	2	Q9IAU4_CHICK	Q9iau4 gallus gall

ALIGNMENTS

RESULT 1

IDA_HUMAN STANDARD; PRT; 391 AA.
 AC Q92838; O75910; Q5JUM7; Q9Y6L0; Q9Y6L1; Q9Y6L2; Q9Y6L3;
 AC Q9Y6L4;
 DT 01-NOV-1997, integrated into UniProtKB/Swiss-Prot.
 DT 15-JUL-1999, sequence version 2.
 DT 07-FEB-2006, entry version 53.
 DE Ectodysplasin-A (Ectodermal dysplasia protein) (EDA protein)
 DE [Contains: Ectodysplasin-A, membrane form; Ectodysplasin-A, secreted form].
 GN Name=EDA; Synonyms=ED1;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
 OC Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [GENOMIC DNA / MRNA] (ISOFORM I), AND VARIANTS EDA HIS-61 AND LEU-69.
 RC TISSUE=Sweat gland;
 RX MEDLINE=98331280; PubMed=8696334;
 RA Kere J., Srivastava A.K., Montonen O., Zonana J., Thomas N.S.T., Ferguson B.M., Munoz F., Morgan D., Clarke A., Baybayan P., Chen E.Y., Ezer S., Saarialho-Kere U., Ia-Chapelle A., Schlessinger D.;
 RT "X-linked anhidrotic (hypohidrotic) ectodermal dysplasia is caused by mutation in a novel transmembrane protein.";
 RL Nat. Genet. 13:409-416(1996).
 RN [2]
 RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM A1), AND VARIANTS EDA. TISSUE=Liver;
 RC MEDLINE=98349961; PubMed=9683615;
 RA Monreal A.W., Zonana J., Ferguson B.M.;
 RT "Identification of a new splice form of the EDA1 gene permits detection of nearly all X-linked hypohidrotic ectodermal dysplasia mutations.";
 RL Am. J. Hum. Genet. 63:380-389(1998).
 RN [3]
 RP NUCLEOTIDE SEQUENCE (ISOFORMS A1; A2; B; C; D; E AND F), AND VARIANTS EDA. MEDLINE=98409495; PubMed=9736768; DOI=10.1093/hmg/7.11.1661;
 RA Bayes M., Hartung A.J., Ezer S., Pispas J., Thesleff I., Srivastava A.K., Kere J.;
 RT "The anhidrotic ectodermal dysplasia gene (EDA) undergoes alternative splicing and encodes ectodysplasin-A with deletion mutations in collagenous repeats.";
 RL Hum. Mol. Genet. 7:1661-1669(1998).
 RN [4]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RX PubMed=15772651; DOI=10.1038/nature03440;
 RA Ross M.T., Grafham D.V., Coffey A.J., Scherer S., McElay K., Muzny D., Platzer M., Howell G.R., Burrows C., Bird C.P., Frankish A., Lovell F.L., Howe K.L., Ashurst J.L., Fulton R.S., Sudbrak R., Jones M.C., Hurler M.E., Andrews T.D., Scott C.E., Searle S.,

RA Rameer J., Whittaker A., Deadman R., Carter N.P., Hunt S.E., Chen R.,
RA Cree A., Gunaratne P., Havlak P., Hodgson A., Metzker M.L.,
RA Richards S., Scott G., Steffen D., Sodergren E., Wheeler D.A.,
RA Worley K.C., Ainscough R., Ambrose K.D., Ansari-Lari M.A., Aradhyia S.,
RA Ashwell R.I., Babarge A.K., Bagguley C.P., Ballabio A., Banerjee R.,
RA Barker G.E., Barlow K.F., Barrett I.P., Bates K.N., Beare D.M.,
RA Beasley H., Beasley O., Beck A., Bethel G., Blechschmidt K., Brady N.,
RA Bray-Allen S., Bridgeman A.M., Brown A.J., Brown M.J., Bonnin D.,
RA Bruford E.A., Buhay C., Burch P., Burford D., Burgess J., Burrill W.,
RA Burton J., Bye J.M., Carder C., Carrel L., Chako J., Chapman J.C.,
RA Chavez D., Chen E., Chen G., Chen Y., Chen Z., Chinnault C.,
RA Ciccodicola A., Clark S.Y., Clarke G., Clee C.M., Clegg S.,
RA Clerc-Blanchet K., Clifford K., Cobley V., Cole C.G., Conquer J.S.,
RA Corby N., Connor R.E., David R., Davies J., Davis C., Davis J.,
RA Delgado O., Deshazo D., Dhani P., Ding Y., Dinh H., Dodsworth S.,
RA Draper H., Dugan-Rocha S., Dunham A., Dunn M., Durbin K.J., Dutta I.,
RA Eades T., Ellwood M., Emery-Cohen A., Errington H., Evans K.L.,
RA Faulkner L., Francis F., Frankland J., Fraser A.E., Gaioczy P.,
RA Gilbert J., Gill R., Gloeckner G., Gregory S.G., Gribble S.,
RA Griffiths C., Grocock R., Gu Y., Gwilliam R., Hamilton C., Hart E.A.,
RA Hawes A., Heath P.D., Heitmann K., Hennig S., Hernandez J.,
RA Hinzmann B., Ho S., Hoffs M., Howden P.J., Huckle E.J., Hume J.,
RA Hunt P.J., Hunt A.R., Isherwood J., Jacob L., Johnson D., Jones S.,
RA de Jong P.J., Joseph S.S., Keenan S., Kelly S., Kershaw J.K., Khan Z.,
RA Kiochis P., Klages S., Knights A.J., Kosiura A., Kovar-Smith C.,
RA Laird G.K., Langford C., Lawlor S., Leversha M., Lewis L., Liu W.,
RA Lloyd C., Lloyd D.M., Loulsegod H., Loveland J.E., Lovell J.D.,
RA Lozado R., Lu J., Lyne R., Ma J., Maheshwari M., Matthews L.H.,
RA McDowell J., McLaren S., McMurray A., Meidl P., Meitinger T.,
RA Milne S., Miner G., Mistry S.L., Morgan M., Morris S., Mueller I.,
RA Mullikin J.C., Nguyen N., Nordieck G., Nyakatura G., O'dell C.N.,
RA Okwuonu G., Palmer S., Pandian R., Parker D., Parrish J.,
RA Pasternak S., Patel D., Pearce A.V., Pearson D.M., Pelan S.E.,
RA Perez L., Porter K.M., Ramsey Y., Reichwald K., Rhodes S.,
RA Ridler K.A., Schlesinger D., Schueler M.G., Sehra H.K.,
RA Shaw-Smith C., Shen H., Sheridan E.M., Showkneen R., Skuce C.D.,
RA Smith M.L., Sotheman H.C., Steingruber H.E., Steward C.A., Storey R.,
RA Swann R.M., Swarbrick D., Taber P.E., Taudien S., Taylor T.,
RA Teague B., Thomas K., Thorpe A., Timms K., Tracey A., Trevanion S.,
RA Tromans A.C., d'Urso M., Verduzco D., Villaseña D., Waldron L.,
RA Wall M., Wang Q., Warren J., Warry G.L., Wei X., West A.,
RA Whitehead S.L., Whiteley M.N., Wilkinson J.E., Willey D.L.,
RA Williams G., Williams L., Williamson A., Williamson H., Wilming L.,
RA Woodmansey R.L., Wray P.W., Yen J., Zhang J., Zhou J., Zoghbi H.,
RA Zorilla S., Buck D., Reinhardt R., Poustka A., Rosenthal A.,
RA Lehtrach H., Meindl A., Minx P.J., Hillier L.W., Willard H.F.,
RA Wilson R.K., Waterston R.H., Rice C.M., Vaudin M., Coulson A.,
RA Nelson D.L., Weinstein G., Sulston J.E., Durbin R., Hubbard T.,
RA Gibbs R.A., Beck S., Rogers J., Bentley D.R.,
RA "The DNA sequence of the human X chromosome.";
RL Nature 434:325-337(2005).
RN [5]
RN TISSUE SPECIFICITY, AND ALTERNATIVE SPLICING.
RP Kobiela K., Kobiela A., Trzciak W.H.;
RT "Expression of a novel transcript isoform of the EDA gene in human
RT umbilical cord.";
RL Eur. J. Hum. Genet. Suppl. 7:104-104(1999).
RN [6]
RN RECEPTOR INTERACTION (ISOFORMS A1 AND A2).
RX MEDLINE=20495245; PubMed=11039935; DOI=10.1126/science.290.5491.523;
RA Yan M., Wang L.-C., Hymowitz S.G., Schillbach S., Lee J., Goddard A.,
RA de Vos A.M., Gao W.-Q., Dixit V.M.;
RT "Two-amino acid molecular switch in an epithelial morphogen that
RT regulates binding to two distinct receptors.";
RL Science 290:523-527(2000).
RN [7]
RN PROCESSING, MUTAGENESIS OF ARG-153, AND CHARACTERIZATION OF VARIANT
RP HIS-156.
RX MEDLINE=21205766; PubMed=11309369; DOI=10.1093/bmg/10.9.953;
RA Elomaa O., Pulkkinen K., Hannelius U., Mikkola M., Saarialho-Kere U.,
RA Kere J.;
RT "Ectodysplasin is released by proteolytic shedding and binds to the
RT EDAR protein.";
RL Hum. Mol. Genet. 10:953-962(2001).
RN [8]
RN CHARACTERIZATION OF VARIANTS CYS-155; CYS-156 AND HIS-156, MUTAGENESIS
RP OF ARG-153; LYS-158 AND ARG-159, AND CLEAVAGE SITE.
RX MEDLINE=21309955; PubMed=11416205; DOI=10.1073/pnas.131076098;
RA Chen Y., Molloy S.S., Thomas L., Gambee J., Baechinger H.P.,
RA Ferguson B.M., Zonana J., Thomas G., Morris N.P.;
RT "Mutations within a furin consensus sequence block proteolytic release
RT of ectodysplasin-A and cause X-linked hypohidrotic ectodermal
RT dysplasia.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:7218-7223(2001).
RN [9]
RN X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 230-391, AND SUBUNIT.
RX PubMed=14656435; DOI=10.1016/j.str.2003.11.009;
RA Hymowitz S.G., Compain D.M., Yan M., Wallweber H.J., Dixit V.M.,
RA Starovasnik M.A., de Vos A.M.;
RT "The crystal structures of EDA-A1 and EDA-A2: splice variants with
RT distinct receptor specificity.";
RL Structure 11:1513-1520(2003).
RN [10]
RN VARIANT EDA TYR-54.
RX MEDLINE=98292028; PubMed=9630076;
RA Hertz J.M., Noergaard Hansen K., Juncker I., Kjeldsen M.,
RA Gregersen N.;
RT "A novel missense mutation (402C-->T) in exon 1 in the EDA gene in a
RT family with X-linked hypohidrotic ectodermal dysplasia.";
RL Clin. Genet. 53:205-209(1998).
RN [11]
RN VARIANT EDA LYS-63.
RX MEDLINE=98168231; PubMed=9507389;
RA Ferguson B.M., Thomas N.S.T., Munoz F., Morgan D., Clarke A.,
RA Zonana J.;
RT "Scarcity of mutations detected in families with X linked hypohidrotic
RT ectodermal dysplasia: diagnostic implications.";
RL J. Med. Genet. 35:112-115(1998).
RN [12]
RN VARIANT EDA ARG-55.
RX MEDLINE=9939307; PubMed=10469321;
RX DOI=10.1046/j.1523-1747.1999.00656.x;
RA Martinez F., Millan J.M., Orellana C., Prieto F.;
RT "X-linked anhidrotic (hypohidrotic) ectodermal dysplasia caused by a
RT novel mutation in EDAL gene: 406T > G (Leu55Arg).";
RL J. Invest. Dermatol. 113:285-286(1999).
RN [13]
RN VARIANTS EDA ARG-60; TYR-252; VAL-269; SER-302 AND MET-378.
RX MEDLINE=21272350; PubMed=11378824; DOI=10.1038/sj.ejhg.5200635;
RA Vincent M.C., Biancalana V., Giniesty D., Mandel J.-L., Calvas P.;
RT "Mutational spectrum of the EDI gene in X-linked hypohidrotic
RT ectodermal dysplasia.";
RL Eur. J. Hum. Genet. 9:355-363(2001).
RN [14]
RN VARIANTS EDA CYS-156; HIS-156; CYS-255; ASP-255; GLY-274; TYR-332 AND
RP THR-349.
RX MEDLINE=21193173; PubMed=11295832; DOI=10.1002/humu.33;
RA Paeeckkoenen K., Cambiaghi S., Novelli G., Ouzts L.V., Penttinen M.,
RA Kere J., Srivastava A.K.;
RT "The mutation spectrum of the EDA gene in X-linked anhidrotic
RT ectodermal dysplasia.";
RL Hum. Mutat. 17:349-349(2001).
CC -!- FUNCTION: Seems to be involved in epithelial-mesenchymal signaling
CC during morphogenesis of ectodermal organs. Isoform A1 binds only
CC to the receptor EDAR, while isoform A2 binds exclusively to the
CC receptor XEDAR.
CC -!- SUBUNIT: Homotrimer. The homotrimers may then dimerize and form
CC higher order oligomers.
CC -!- INTERACTION:
CC Q9UNE0:EDAR; NbExp=1; IntAct=EBI-529425, EBI-529289;
CC -!- SUBCELLULAR LOCATION: Type II membrane protein and secreted.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=8;
CC Comment=Additional isoforms seem to exist;
CC Name=A1; Synonyms=II;
CC IsoId=Q92838-1; Sequence=Displayed;

CC	Name=I;	CC	-I- PTM: N-glycosylated (By similarity).
CC	IsolId=Q92838-2; Sequence=VSP_006454, VSP_006455;	CC	-I- PTM: Processing by furin produces a secreted form (By similarity).
		CC	-I- DISEASE: Defects in EDA are the cause of anhidrotic ectodermal dysplasia. The disease is characterized by sparse hair (atrachosis or hypotrichosis), abnormal or missing teeth and the inability to sweat due to the absence of sweat glands.
Qy	Query Match 100.0%; Score 2098; DB 1; Length 391;	CC	-I- SIMILARITY: Belongs to the tumor necrosis factor family.
Db	Best Local Similarity 100.0%; Pred. No. 3.3e-121;	CC	-I- SIMILARITY: Contains 1 collagen-like domain.
	Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	CC	-----
		CC	Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
		CC	Distributed under the Creative Commons Attribution-NoDerivs License
		CC	-----
Qy	1 MGYPEVERRELLPAAAPRRGSGCGCGGAPARAGEGNSCLLFLGFFGLSLALHLTLCC 60	CC	EMBL; AJ300468; CAC29151.1; -; Genomic_DNA.
Db	1 MGYPEVERRELLPAAAPRRGSGCGCGGAPARAGEGNSCLLFLGFFGLSLALHLTLCC 60	DR	EMBL; AJ300469; CAC29151.1; JOINED; Genomic_DNA.
		DR	EMBL; AJ278907; CAC29151.1; JOINED; Genomic_DNA.
Qy	61 YLELSELRRERGAESRLGSGTPTSGTLSSLGGLDPDPSITSHLGQSPKQOPLPEGE 120	DR	EMBL; AJ300468; CAC29152.1; -; Genomic_DNA.
Db	61 YLELSELRRERGAESRLGSGTPTSGTLSSLGGLDPDPSITSHLGQSPKQOPLPEGE 120	DR	EMBL; AJ300469; CAC29152.1; JOINED; Genomic_DNA.
		DR	EMBL; AJ278907; CAC29152.1; JOINED; Genomic_DNA.
Qy	121 AALHSDSQDGHQWALLNFFPDPEKPYSEESRRVRNKRKSKNEGADGVFNKKGKAG 180	DR	SMR; Q9BEG5; 242-390.
Db	121 AALHSDSQDGHQWALLNFFPDPEKPYSEESRRVRNKRKSKNEGADGVFNKKGKAG 180	DR	InterPro; IPR008160; Collagen.
		DR	InterPro; IPR006052; TNF family.
Qy	181 PRGPNPGRPDPGPPGQGGPIGPIGTGTTVMGPPGPPGPPGPPGPPGPPGPPGPPG 240	DR	Pfam; PF01391; Collagen; 1.
Db	181 PRGPNPGRPDPGPPGQGGPIGPIGTGTTVMGPPGPPGPPGPPGPPGPPGPPGPPG 240	DR	Pfam; PF00229; TNF; 1.
		DR	PROSITE; PS00251; TNF_1; FALSE_NEG.
Qy	241 AGTRENQPAVHLQGGSAIQVKNDSLGGVLDWDSRITWNPVKVFLHPRSGEVLVDGT 300	DR	PROSITE; PS00449; TNF_2; 1.
Db	241 AGTRENQPAVHLQGGSAIQVKNDSLGGVLDWDSRITWNPVKVFLHPRSGEVLVDGT 300	KW	Alternative splicing; Collagen; Membrane; Signal-anchor; Transmembrane.
		FT	Differentiation; Glycoprotein; Ectodysplasin-A, membrane form.
Qy	301 YFIYSQVEVYIINFDTFASVEVVDVKPFLOQTRSIETGKTNNTCYTAGVCLLKARQKI 360	CHAIN	1 391
Db	301 YFIYSQVEVYIINFDTFASVEVVDVKPFLOQTRSIETGKTNNTCYTAGVCLLKARQKI 360	FT	Ectodysplasin-A, secreted form (By similarity).
		FT	/FTId=PRO_0000034536.
Qy	361 AVKMVHADISINMSKHTTFFGAIRLGEAPAS 391	FT	Ectodysplasin-A, secreted form (By similarity).
Db	361 AVKMVHADISINMSKHTTFFGAIRLGEAPAS 391	FT	/FTId=PRO_0000034537.
		FT	Cytoplasmic (Potential).
		FT	Signal-anchor for type II membrane protein (Potential).
		FT	Extracellular (Potential).
		FT	Collagen-like.
		FT	Cleavage (by furin) (By similarity).
		FT	N-linked (GlcNAc. . .) (Potential).
		FT	N-linked (GlcNAc. . .) (Potential).
		FT	Missing (in isoform A2).
		FT	/FTId=VSP_006453.
		SQ	SEQUENCE 391 AA; 41567 MW; 1F87AD67A04EB7AA CRC64;
			Query Match 94.0%; Score 1972; DB 1; Length 391;
			Best Local Similarity 94.4%; Pred. No. 1.9e-113;
			Matches 369; Conservative 4; Mismatches 18; Indels 0; Gaps 0;
Qy	1 MGYPEVERRELLPAAAPRRGSGCGCGGAPARAGEGNSCLLFLGFFGLSLALHLTLCC 60	Qy	1 MGYPEVERRELLPAAAPRRGSGCGCGGAPARAGEGNSCLLFLGFFGLSLALHLTLCC 60
Db	1 MGYPEVERRELLPAAAPRRGSGCGCGGAPARAGEGNSCLLFLGFFGLSLALHLTLCC 60	Db	1 MGYPEVERRELLPAAAPRRGSGCGCGGAPARAGEGNSCLLFLGFFGLSLALHLTLCC 60
		Qy	61 YLELSELRRERGAESRLGSGTPTSGTLSSLGGLDPDPSITSHLGQSPKQOPLPEGE 120
		Db	61 YLELSELRRERGAESRLGSGTPTSGTLSSLGGLDPDPSITSHLGQSPKQOPLPEGE 120
		Qy	121 AALHSDSQDGHQWALLNFFPDPEKPYSEESRRVRNKRKSKNEGADGVFNKKGKAG 180
		Db	121 AALHSDSQDGHQWALLNFFPDPEKPYSEESRRVRNKRKSKNEGADGVFNKKGKAG 180
		Qy	181 PRGPNPGRPDPGPPGQGGPIGPIGTGTTVMGPPGPPGPPGPPGPPGPPGPPGPPG 240
		Db	181 PRGPNPGRPDPGPPGQGGPIGPIGTGTTVMGPPGPPGPPGPPGPPGPPGPPGPPG 240
		Qy	241 AGTRENQPAVHLQGGSAIQVKNDSLGGVLDWDSRITWNPVKVFLHPRSGEVLVDGT 300
		Db	241 AGTRENQPAVHLQGGSAIQVKNDSLGGVLDWDSRITWNPVKVFLHPRSGEVLVDGT 300
		Qy	301 YFIYSQVEVYIINFDTFASVEVVDVKPFLOQTRSIETGKTNNTCYTAGVCLLKARQKI 360
		Db	301 YFIYSQVEVYIINFDTFASVEVVDVKPFLOQTRSIETGKTNNTCYTAGVCLLKARQKI 360
		Qy	361 AVKMVHADISINMSKHTTFFGAIRLGEAPAS 391
		Db	361 AVKMVHADISINMSKHTTFFGAIRLGEAPAS 391

			RESULT 2
			ED_ BOVIN
			ID_ EDA_ BOVIN
			AC Q9BEG5; Q9BEG6; STANDARD; PRT; 391 AA.
			DT 06-JUN-2002, integrated into UniProtKB/Swiss-Prot.
			DT 06-JUN-2002, sequence version 2.
			DT 07-FEB-2006, entry version 30.
			DE Ectodysplasin-A (Ectodysplasin-1) (Ectodermal dysplasia protein)
			DE [Contains: Ectodysplasin-A, membrane form; Ectodysplasin-A, secreted form].
			GN Bos taurus (Bovine);
			OS Bos taurus (Bovine);
			OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
			OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
			OC Pecora; Bovidae; Bovinae; Bos.
			OX NCBI_TaxID=9913;
			RN [1]
			RP NUCLEOTIDE SEQUENCE (GENOMIC DNA) (ISOFORMS A1 AND A2).
			RC STRAIN=Holstein;
			RX MEDLINE=21070494; PubMed=11167539;
			RA Drogemuehler C., Distl O., Leeb T.;
			RT "Identification of a highly polymorphic microsatellite within the
			RT bovine ectodysplasin A (ED1) gene on BTA Xq22-24.";
			RL Anim. Genet. 31:416-416(2000).
			CC -I- FUNCTION: Probably involved in epithelial-mesenchymal signaling.
			CC Isoform A1 binds only to the receptor EDAR, while isoform A2 binds
			CC exclusively to the receptor XEDAR (By similarity).
			CC -I- SUBUNIT: Homotrimer (By similarity).
			CC -I- SUBCELLULAR LOCATION: Type II membrane protein and secreted (By
			CC similarity).
			CC -I- ALTERNATIVE PRODUCTS:
			CC Event=Alternative splicing; Named isoforms=2;
			CC Comment=Additional isoforms seem to exist;
			CC Name=A1;
			CC IsolId=Q9BEG5-1; Sequence=Displayed;
			CC Name=A2;
			CC IsolId=Q9BEG5-2; Sequence=VSP_006453;

SQ SEQUENCE 386 AA; 41235 MW; 6095BDBC3EACF68B CRC64;

Query Match 90.1%; Score 1890.5; DB 2; Length 386;
Best Local Similarity 91.3%; Pred. No. 1.9e-108;
Matches 357; Conservative 8; Mismatches 21; Indels 5; Gaps 2;

QY 1 MGYPEVERRELLPAAAPRRGSGCGCGGAPARAGSGNCLLLFLGFLSLALHLTLCC 60
DB 1 MGYPEVERREPLTAAPRRGSGCGCGGAPARAGSGNCLLLFLGFLSLALHLTLCC 60
QY 61 YLELRSELRRERGAERLGGSGTGTGSLSSGLGLDPPSPITSHLGQSPKQOPLPGE 120
DB 61 YLELRSELRRERGAERLGGSGTGTGSLSSGLGLDPPSPITSHLGQSPKQOPLPGE 120
QY 121 AALHSQDQHOMALLNFFPDDEKPYSEESRRVRNKKSKSNEGADGVPVNNKKKKKAG 180
DB 117 TALPHSRDQHOMALLNFFPEESKSYSEDE-RRFRNRKRSKSEGTGDPVNNKKKKKAG 175
QY 181 PGPNGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 240
DB 176 PGPNGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 235
QY 241 AGPRENQPAVHLQGGSAIQVKNLSSGVLDWSRITMNPVKFKLHPRSGEVLVDGT 300
DB 236 AGPRENQPAVHLQGGSAIQVKNLSSGVLDWSRITMNPVKFKLHPRSGEVLVDGT 295
QY 301 YFTYSQVEVYINFTDPSYEVVVDKPFLOCTRSIETGKNTYNTCYTAGVCLLKARQKI 360
DB 296 YFTYSQVEVYINFTDPSYEVVVDKPFLOCTRSIETGKNTYNTCYTAGVCLLKARQKI 355
QY 361 AVKQVHADISINMSKHTTFFGAILRGEAPAS 391
DB 356 AVKQVHADISINMSKHTTFFGAILRGEAPAS 386

RESULT 6
Q5EFZ5_CHICK PRELIMINARY; PRT; 356 AA.

ID Q5EFZ5_CHICK
AC Q5EFZ5;
DT 15-MAR-2005, integrated into UniProtKB/TREMBL.
DT 15-MAR-2005, sequence version 1.
DT 07-FEB-2006, entry version 6.
DE Ectodysplasin A1 (Fragment).
GN Name=Eda;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Houghton L., Morgan B.A.;
RL Submitted (JAN-2005) to the EMBL/GenBank/DBJ databases.
CC
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs license
CC
CC EMBL; AY885699; AAW81692.1; -; mRNA.
DR SMR; O5EFZ5; 207-355.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0016020; F:tumor necrosis factor receptor binding; IEA.
DR GO; GO:0005164; F:tumor necrosis factor response; IEA.
DR GO; GO:0006955; P:immune response; IEA.
DR GO; GO:006817; P:phosphate transport; IEA.
DR InterPro; IPR008160; Collagen.
DR InterPro; IPR006052; TNF_family.
DR Pfam; PF01391; Collagen; 2.
DR Pfam; PF00229; TNF; 1.
DR PROSITE; PS50049; TNF_2; 1.
KW Collagen.
FT NON_TER
SQ SEQUENCE 356 AA; 37527 MW; 8D53FD97F71FDDFB CRC64;

Query Match 68.7%; Score 1441.5; DB 2; Length 356;
Best Local Similarity 74.9%; Pred. No. 7.7e-81;
Matches 289; Conservative 19; Mismatches 41; Indels 37; Gaps 9;

QY 13 PAAAPRRGSG---CGCGGAPARAGSGNCLLLFLGFLSLALHLTLCCYLELRSEL 68
DB 1 PAAAGARRGEAARGGSCAG-----GSMFLFLGFFALSALAHVLTLCYLELRSEL 53
QY 69 RREGAERLGGSGTGTGSLSSGLGLDPPSPITSHLGQSPK-QOPLPGEAALHSDS 127
DB 54 RRDGFGPQ---AAPRRDGTAAAA---PGAP-----PAVRPQPAESGERR----- 94
QY 128 QDQHOMALLNFFPDDEKPYSEESRRVRNKKSKSNEGADGVPVNNKKKKKAGPPGP 185
DB 95 ---QQLALLNFFPEEKJH-VSEGRVRNKKSKSGEGDPGSSVNNKKKKKAGPPGP 150
QY 186 GPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 245
DB 151 GQGGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 210
QY 246 NPQAVVHLQGGSAIQVKNLSSGVLDWSRITMNPVKFKLHPRSGEVLVDGTFTYIS 305
DB 211 TQPAVHLQGGSAIQVKNLSSGVLDWSRITMNPVKFKLHPRSGEVLVDGTFTYIS 270
QY 306 QVEVYINFTDPSYEVVVDKPFLOCTRSIETGKNTYNTCYTAGVCLLKARQKI 365
DB 271 QVEVYINFTDPSYEVVVDKPFLOCTRSIETGKNTYNTCYTAGVCLLKARQKI 330
QY 366 HADISINMSKHTTFFGAILRGEAPAS 391
DB 331 HADISINMSKHTTFFGAILRGEAPAS 356

RESULT 7
Q50D54_GASAC PRELIMINARY; PRT; 367 AA.

ID Q50D54_GASAC
AC Q50D54;
DT 07-JUN-2005, integrated into UniProtKB/TREMBL.
DT 07-JUN-2005, sequence version 1.
DT 07-FEB-2006, entry version 6.
DE Ectodysplasin.
GN Name=Eda;
OS Gasterosteus aculeatus (Three-spined stickleback).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Gasterosteiformes;
OC Gasterosteidae; Gasterosteus.
OX NCBI_TaxID=69293;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15790847; DOI=10.1126/science.1107239;
RA Colosimo P.F., Hosemann K.E., Balabhadra S., Villarreal G. Jr.,
RA Dickson M., Hightower J., Schmutz J., Myers R.M., Schluter D.,
RA Kingsley D.M.;
RT Widespread parallel evolution in sticklebacks by repeated fixation of
RT Ectodysplasin alleles."
RL Science 307:1928-1933(2005).
CC
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs license
CC
CC EMBL; AY897589; AAY27076.1; -; Genomic_DNA.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005164; F:tumor necrosis factor receptor binding; IEA.
DR GO; GO:0006955; P:immune response; IEA.
DR GO; GO:0006817; P:phosphate transport; IEA.
DR InterPro; IPR008160; Collagen.
DR InterPro; IPR006052; TNF_family.
DR Pfam; PF01391; Collagen; 1.
DR Pfam; PF00229; TNF; 1.
DR PROSITE; PS00351; TNF_1; UNKNOWN_1.
DR PROSITE; PS00351; TNF_2; 1.


```

RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
CC -----
CC Copyrighted by the Uniprot Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NoDerivs License
CC -----
DR EMBL; BC013626; AAH13626.1; -; mRNA.
DR Ensembl; ENSMUSG0000026043; Mus musculus.
DR MGI; 88453; Col3a1.
DR GO; GO:0005737; Cytoplasm; IEA.
DR GO; GO:0006817; P; phosphate transport; IEA.
DR InterPro; IPR008161; Clg helix.
DR InterPro; IPR008160; Collagen.
DR Pfam; PF01391; Collagen; 12.
DR ProDom; PD000007; Clg_helix; 2.
KW Collagen.
FT NON_TER
SQ SEQUENCE 998 AA; 95451 MW; 2AB4BA953B7084A6 CRC64;
Query Match 13.5%; Score 284; DB 2; Length 998;
Best Local Similarity 30.3%; Pred. No. 4e-09;
Matches 87; Conservative 15; Mismatches 91; Indels 94; Gaps 11;
QY 13 PAAAPRERSQG-----CGCGGAPARAGEGNSCLLFLGFFGLSLALHLTLCCYLELRSEL 68
DB 548 PRGGGGERGEGHGGPPGAPGAPGAPGAPGAPGAPGAPGAPGAPGAPGAPGAPGAPGAPGAP 578
QY 69 RRRGAEERLGGSGTGTGTSGLSLGLDPDPSITSHLGSPSPKQOPLSPGERALHSDSQ 128
DB 579 KGERGAPGKGEGGPPGAPGAPGAPGAPGAPGAPGAPGAPGAPGAPGAPGAPGAPGAPGAP 628
QY 129 DGHQALLNFFFPDEKPYSEESRRVRNRKSKNSGADGPPVKNKKKGKAGPPGPPGPP 188
DB 629 -----FP-----GGRLGPPGPN-----NGNPPGPPGSGAP 654
QY 189 GPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP 245
DB 655 GKDPGPPGAGNSGPNPGIAGPKGAGQPGKPPGAPGAPGAPGAPGAPGAPGAPGAPGAP 714
QY 246 NQPAV-----VHLOQGSATQVND-----LSGGVLDMSRITWNPVVF 284
DB 715 GPPGPPGPPRVALDLRVRVKVQEPVIMGNV-----VLLDPKVF 755
RESULT 12
Q8C979_MOUSE PRELIMINARY; PRT; 742 AA.
AC Q8C979;
DT 01-MAR-2003, integrated into UniProtKB/TrEMBL.
DT 01-MAR-2003, sequence version 1.
DE 07-FEB-2006, entry version 23.
DE 7 days neonate cerebellum cDNA, RIKEN full-length enriched library,
DE clone: A730023E20 product: collectin sub-family member 12, full insert
DE sequence.
GN Name=Collect2;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX MEDLINE=99279233; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RL "High-efficiency full-length cDNA cloning."
RL Methods Enzymol. 303:19-44(1999).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX PubMed=16141072; DOI=10.1126/science.1112014;
RA Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N.,
RA Oyama R., Ravasi T., Lenhard B., Wells K., Kodzius R., Shimokawa K.,
RA Bajic V.B., Brenner S.E., Batalov S., Forrest A.R., Zavolan M.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,

```

```

RA Davis M.J., Wilming L.G., Aldinis V., Allen J.E.,
RA Ambei-Imbimbato A., Aweiler R., Aturaliya R.N., Bailey T.L.,
RA Banaal M., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,
RA Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R.,
RA Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,
RA di Bernardo D., Down T., Engstrom P., Fagioli M., Faulkner G.,
RA Fletcher C.F., Fukushima T., Furuno M., Futaki S., Gariboldi M.,
RA Georgii-Hemming P., Gingeras T.R., Gojobori T., Green R.E.,
RA Gustincich S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N.,
RA Hill D., Huminicki L., Iacono M., Ikeo K., Iwama A., Ishikawa T.,
RA Jakt M., Kanapin A., Katoh M., Kawasawa Y., Kelso J., Kitamura H.,
RA Kitano H., Kollas G., Krishnan S.P., Kruger A., Kummerfeld S.K.,
RA Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J.,
RA Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,
RA Matsuoka H., Matsuura S., Miki H., Mignone F., Miyake S., Morris K.,
RA Mottagui-Tabar S., Mulder N., Nakano N., Nakachi H., Ng P.,
RA Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,
RA Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavoni G., Pesole G.,
RA Petrovsky N., Piazza S., Reed J., Reid J.F., Ring B.Z., Ringwald M.,
RA Rost B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C.,
RA Schonbach C., Sekiguchi K., Semple C.A., Seno S., Sessa L., Sheng Y.,
RA Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,
RA Sperling S., Stupka E., Sultana K., Takenaka Y., Taki K.,
RA Tammoja K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,
RA Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yagi K.,
RA Yamanishi H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C.,
RA Grimmond S.M., Teasdale R.D., Liu E.T., Brusic V., Quackenbush J.,
RA Wahlestedt C., Mattick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,
RA Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T.,
RA Iida J., Imamura K., Itoh T., Kato T., Kawaji H., Kawagashira N.,
RA Kawashima T., Kojima M., Kondo S., Konno H., Nakano K., Ninomiya N.,
RA Nishio T., Okada M., Plessey C., Shibata K., Shiraki T., Suzuki S.,
RA Tagami M., Waki K., Watahiki A., Okamura-Oho Y., Suzuki H., Kawai J.,
RA Hayashizaki Y.;
RT "The transcriptional landscape of the mammalian genome."
RL Science 309:1559-1563(2005).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX PubMed=16141073; DOI=10.1126/science.1112009;
RG RIKEN Genome Exploration Research Group, and Genome Science Group
RG (Genome Network Core Team) and the FANTOM Consortium;
RT "Antisense Transcription in the Mammalian Transcriptome."
RL Science 309:1564-1566(2005).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaide I., Oosato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran M.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wymshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hironaka-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shingawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,

```


DB	969	GHPGPPGPEQGLPGLAGECTK-GDPGAGLPGK	1003
RESULT 15			
COSAL HUMAN			
ID	COSAL HUMAN	STANDARD;	PRT; 1838 AA.
AC	203008; Q15094; Q5UX4;		
DT	01-FEB-1991,	integrated into UniProtKB/Swiss-Prot.	
DT	13-SEP-2005,	sequence version 3.	
DT	07-MAR-2006,	entry version 71.	
DE	Collagen alpha-1(V) chain precursor.		
GN	Name=COL5A1;		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;		
OC	Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]_		
RP	NUCLEOTIDE SEQUENCE [MRNA], AND PROTEIN SEQUENCE OF 556-565.		
RX	MEDLINE=91302336; PubMed=2071595;		
RA	Takahara K., Seto Y., Okasawa K., Okamoto N., Noda A., Yaoi Y.,		
RA	Kato I.;		
RT	"Complete primary structure of human collagen alpha 1 (V) chain.;"		
RL	J. Biol. Chem. 266:13124-13129(1991).		
RN	[2]_		
RP	NUCLEOTIDE SEQUENCE [MRNA].		
RX	MEDLINE=92105142; PubMed=1722213;		
RA	Greenpan D.S., Cheng W., Hoffman G.G.;		
RT	"The pro-alpha-1(V) collagen chain: complete primary structure,		
RT	distribution of expression, and comparison with the pro-alpha-1(XI)		
RT	collagen chain.;"		
RL	J. Biol. Chem. 266:24727-24733(1991).		
RN	[3]_		
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].		
RX	PubMed=15164053; DOI=10.1038/nature02465;		
RA	Humphray S.J., Oliver K., Hunt A.R., Plumb R.W., Loveland J.E.,		
RA	Howe K.L., Andrews T.D., Searle S., Hunt S.E., Scott C.E., Jones M.C.,		
RA	Ainscough R., Almeida J.P., Ambrose K.D., Ashwell R.I.S.,		
RA	Babbage A.K., Babbage S., Bagguely C.L., Bailey J., Banerjee R.,		
RA	Barker D.J., Barlow K.F., Bates K., Beasley H., Beasley O., Bird C.P.,		
RA	Bray-Allen S., Brown A.J., Brown J.Y., Burford D., Burtill W.,		
RA	Burton J., Carder C., Carter N.P., Chapman J.C., Chen Y., Clarke G.,		
RA	Clark S.Y., Clee C.M., Clegg S., Collier R.E., Corby N., Crosier M.,		
RA	Cummings A.T., Davies J., Dhali P., Dunn M., Dutta I., Dyer L.W.,		
RA	Frankland J.A., Faulkner L., Fleming C.J., Frankish A.,		
RA	Frankland J.A., French L., Fricker D.G., Garner P., Garnett J.,		
RA	Ghorji J., Gilbert J.G.R., Glison C., Grafham D.V., Gribble S.,		
RA	Griffiths C., Griffiths-Jones S., Grocock R., Guy J., Hall R.E.,		
RA	Hammond S., Harley J.L., Harrison E.S.I., Hart E.A., Heath P.D.,		
RA	Henderson C.D., Hopkins B.L., Howard P.J., Howden P.J., Huckle E.,		
RA	Johnson C., Johnson D., Joy A.A., Kay M., Keenan S., Kershaw J.K.,		
RA	Kimberley A.M., King A., Knights A., Laird G.K., Langford C.,		
RA	Lawlor S., Leongamornlert D.A., Leversha M., Lloyd C., Lloyd D.M.,		
RA	Lovell J., Martin S., Mashreghi-Mohammadi M., Matthews L., McLaren S.,		
RA	McLay K.E., McMurray A., Milne S., Nickerson T., Nisbett J.,		
RA	Nordsiek G., Pearce A.V., Peck A.I., Porter K.M., Pandian R.,		
RA	Pelan S., Phillimore B., Povey S., Ramsey Y., Rand V., Scharfe M.,		
RA	Sandra H.K., Showkeen R., Sims S.K., Skuce C.D., Smith M.,		
RA	Steward C.A., Swarbreck D., Sycamore N., Tester J., Thorpe A.,		
RA	Tracey A., Tromans A., Thomas D.W., Wall M., Wallis J.M., West A.P.,		
RA	Whitehead S.L., Willey D.L., Williams S.A., Wilming L., Wray P.W.,		
RA	Young L., Ashurst J.L., Coulson A., Blocker H., Durbin R.,		
RA	Sulston J.E., Hubbard T., Jackson M.J., Bentley D.R., Beck S.,		
RT	Rogers J., Dunham I.;		
RT	"DNA sequence and analysis of human chromosome 9.;"		
RL	Nature 429:369-374(2004).		
RN	[4]_		
RP	PROTEIN SEQUENCE OF 621-822.		
RC	TISSUE=Chorioamniotic membrane;		
RX	MEDLINE=89227189; PubMed=2496661;		
RA	Seyer J.M., Kang A.H.;		
RT	"Covalent structure of collagen: amino acid sequence of three cyanogen		
RT	bromide-derived peptides from human alpha 1(V) collagen chain.;"		
RL			
RL	Arch. Biochem. Biophys. 271:120-129(1989).		
RN	[5]_		
RP	PROTEIN SEQUENCE OF 823-950, AND HEPARIN-BINDING.		
RX	MEDLINE=90366601; PubMed=2203476; DOI=10.1016/0304-4165(90)90108-9;		
RA	Yaoi Y., Hashimoto K., Koitabashi H., Takahara K., Ito M., Kato I.;		
RT	"Primary structure of the heparin-binding site of type V collagen.;"		
RL	Biochim. Biophys. Acta 1035:139-145(1990).		
RN	[6]_		
RP	PROTEIN SEQUENCE OF 556-571.		
RC	TISSUE=Placenta;		
RX	MEDLINE=92239022; PubMed=15711108;		
RA	Mann K.;		
RT	"Isolation of the alpha 3-chain of human type V collagen and		
RT	characterization by partial sequencing.;"		
RL	Biol. Chem. Hoppe-Seyler 373:69-75(1992).		
RN	[7]_		
RP	PROTEIN SEQUENCE OF 565-576; 756-772; 1012-1029; 1219-1232 AND		
RP	1465-1477.		
RC	TISSUE=Chorioamniotic membrane;		
RX	MEDLINE=94337164; PubMed=8181482;		
RA	Moradi-Ameli M., Rousseau J.C., Klemm J.P., Champlaud M.F.,		
RA	Boutillon M.M., Bernillon J., Wallach J.M., van der Rest M.;		
RT	"Diversity in the processing events at the N-terminus of type-V		
RT	collagen.;"		
RL	Eur. J. Biochem. 221:987-995(1994).		
RN	[8]_		
RP	INTERACTION WITH CSPG4.		
RX	PubMed=9099729; DOI=10.1074/jbc.272.16.10769;		
RA	Tillet E., Ruggiero F., Nishiyama A., Stallcup W.B.;		
RT	"The membrane-spanning proteoglycan NG2 binds to collagens V and VI		
RT	through the central nonglobular domain of its core protein.;"		
RL	J. Biol. Chem. 272:10769-10776(1997).		
RN	[9]_		
RP	DISEASE, AND VARIANT EDS-I SER-1639.		
RX	MEDLINE=97195540; PubMed=9042913;		
RA	de Paeppe A., Nuytinck L., Hausser I., Anton-Lamprecht I.,		
RA	Naeyaert J.-M.;		
RT	"Mutations in the COL5A1 gene are causal in the Ehlers-Danlos		
RT	syndromes I and II.;"		
RL	Am. J. Hum. Genet. 60:547-554(1997).		
RN	[10]_		
RP	VARIANTS EDS-1 SER-530 AND ASP-1489.		
RX	MEDLINE=20068401; PubMed=10602121;		
RA	DOI=10.1002/(SICI)1096-8628(20000103)90:1<72::AID-AJMG13>3.0.CO;2-C;		
RA	Giunta C., Steinmann B.;		
RT	"Compound heterozygosity for a disease-causing G1489D and disease-		
RT	modifying G530S substitution in COL5A1 of a patient with the classical		
RT	type of Ehlers-Danlos syndrome: an explanation of intrafamilial		
RT	variability?;"		
RL	Am. J. Med. Genet. 90:72-79(2000).		
CC	-I- FUNCTION: Type V collagen is a member of group I collagen		
CC	(fibrillar forming collagen). It is a minor connective tissue		
CC	component of nearly ubiquitous distribution. Type V collagen binds		
CC	to DNA, heparan sulfate, thrombospondin, heparin, and insulin.		
CC	-I- SUBUNIT: Trimers of two alpha 1(V) and one alpha 2(V) chains in		
CC	most tissues and trimers of one alpha 1(V), one alpha 2(V), and		
CC	one alpha 3(V) chains in placenta. Interacts with CSPG4.		
CC	-I- PTM: Prolines at the third position of the tripeptide repeating		
CC	unit (G-X-Y) are hydroxylated in some or all of the chains.		
CC	-I- PTM: Sulfated on 40% of tyrosines.		
CC	-I- DISEASE: Defects in COL5A1 are a cause of Ehlers-Danlos syndrome		
CC	type I (EDS-I) [MIM:130000]; also known as Ehlers-Danlos syndrome		
CC	gravis. EDS-I is a connective-tissue disorder characterized by		
CC	loose-jointedness and fragile, velvety, stretchable, bruisable		
CC	skin that heals with peculiar 'cigarette-paper' scars. Inheritance		
CC	is autosomal dominant		
CC	-I- DISEASE: Defects in COL5A1 are a cause of Ehlers-Danlos syndrome		
CC	type II (EDS-II) [MIM:130010]; also known as Ehlers-Danlos		
CC	syndrome mitis. Inheritance is autosomal dominant.		
CC	-I- SIMILARITY: Belongs to the fibrillar collagen family.		
CC	-I- SIMILARITY: Contains 1 laminin G-like domain.		
CC	-I- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.		
CC	-----		

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NoDerivs License

CC EMBL; D90279; BAA14323.1; -; mRNA.
CC EMBL; M76729; AAAS9993.1; -; mRNA.
DR EMBL; AL591890; CAI15483.1; -; Genomic DNA.
DR EMBL; AL603650; CAI15483.1; JOINED; Genomic DNA.
DR EMBL; AL645768; CAI15483.1; JOINED; Genomic DNA.
DR EMBL; AL645768; CAI17261.1; -; Genomic DNA.
DR EMBL; AL591890; CAI17261.1; JOINED; Genomic DNA.
DR EMBL; AL603650; CAI17261.1; JOINED; Genomic DNA.
DR EMBL; AL603650; CAI39859.1; -; Genomic DNA.
DR EMBL; AL591890; CAI39859.1; JOINED; Genomic DNA.
DR EMBL; AL645768; CAI39859.1; JOINED; Genomic DNA.
DR PIR; S18802; CGHUIV.
DR PDB; 1A89; Model; A/B/C=904-924.
DR PDB; 1A9A; Model; A/C=904-924.
DR Ensembl; ENSG00000130635; Homo sapiens.
DR HGNC; HGNC:2209; COL5A1.
DR MIM; 120215; Gene.
DR MIM; 130000; Phenotype.
DR MIM; 130010; Phenotype.
DR GO; GO:0005588; C:collagen type V; TAS.
DR InterPro; IPR008161; Clg_helix.
DR InterPro; IPR008160; Collagen.
DR InterPro; IPR000885; Fib_collagen_C.
DR InterPro; IPR001791; Laminin_G.
DR InterPro; IPR012680; Laminin_G_2.
DR InterPro; IPR003129; Laminin_G_TSP_N.
DR Pfam; PF01410; COLFI; 1.
DR Pfam; PF01391; Collagen; 17.
DR Pfam; PF02210; Laminin_G_2; 1.
DR ProDom; PD000007; Clg_helix; 4.
DR ProDom; PD002078; Fib_collagen_C; 1.
DR SMART; SM00038; COLFI; 1.
DR SMART; SM00282; LamG; 1.
DR SMART; SM00210; TSPN; 1.
DR PROSITE; PS50025; LAM_G_DOMAIN; FALSE NEG.
KW 3D-structure; Collagen; Direct protein sequencing; Disease mutation;
KW Ehlers-Danlos syndrome; Extracellular matrix; Heparin-binding;
KW Hydroxylation; Repeat; Signal; Structural protein; Sulfation.
FT SIGNAL 1 37 Potential.
FT CHAIN 38 1605 Collagen alpha-1(V) chain.
FT PROPEP 1606 1838 /FTid=PRO_0000005756.
FT DOMAIN 39 230 Laminin G-like.
FT REGION 72 244 Nonhelical region.
FT REGION 231 443 Interrupted collagenous region.
FT REGION 444 558 Triple-helical region.
FT REGION 559 1570 Nonhelical region.
FT REGION 1571 1605 Sulfotyrosine (Potential).
FT MOD_RES 234 234 Sulfotyrosine (Potential).
FT MOD_RES 236 236 Sulfotyrosine (Potential).
FT MOD_RES 240 240 Sulfotyrosine (Potential).
FT MOD_RES 262 262 Sulfotyrosine (Potential).
FT MOD_RES 263 263 Sulfotyrosine (Potential).
FT MOD_RES 338 338 Sulfotyrosine (Potential).
FT MOD_RES 340 340 Sulfotyrosine (Potential).

Query Match 13.4%; Score 280.5; DB 1; Length 1838;
Best Local Similarity 30.1%; Pred. No. 1.2e-08;
Matches 83; Conservative 9; Mismatches 69; Indels 115; Gaps 10;
QY 13 PAAAPRERGQCG-GCGGAPARAGCNSCLFLGFFGLSLALHLLTLCCYLELRSELRE 71
Db 857 PLGPPGEGKLGVPGLPGYGRGPKGS-----IGFPG?P----- 891
QY 72 RGNESRLGSGTGTGTSGLSSGLGLDPDPSITSHLCQSPKQPLEPGEAALHSDSQGH 131
Db 892 -GANGKGGRGTP-----GKPGRGQGTG----- 916
QY 132 QMALLNFFFPDEKPYSEESRRVRNRKSKNEGADGPVKNKKKGKAGPP---GPNGPP 188

Db 917 -----PRGERGPRGITKGPCKGNSGGDGP-----AGPPGERGPNGPQ 954
QY 189 GPPGPPGQPPGIPGIPGIPG-----TTVMGPPGPPGP----- 222
Db 955 GPTGPPGKPPGPPGKDGGLPGHPGQGETGFGKGTGPPGPPGVVGPQGTGTGPMGR 1014
QY 223 --PGPQPPGLQGPSSGAADKAGTRENQPAVVHLQQQ 256
Db 1015 GHPPGPPGEGGLFGLAGKCTK-GDPCPAGLPGK 1049

Search completed: June 2, 2006, 17:30:03
Job time : 301 secs

QY 1 MGYPERRELLPAAAPRRSGCGCGAPARAGSGNCLLFLGFFGLSLALHLTLCC 60
DB 1 MGYPERREPLPAAAPRRSGCGCGCRGAPARAGSGNCLLFLGFFGLSLALHLTLCC 60
QY 61 YLELSELRRERGAESRLGSGTPTGTLSSLGGLDPSPTITSHLGQSPKQOPLPGE 120
DB 61 YLELSELRRERGTESRLGSGPAGTSGTLSSPGSLDPVGPITRHLGQSPKQOPLPGE 120
QY 121 AALHSDSQDHQWALLNFFPDEKPYSEESRRVRNRKSKSNEGADGPNKKNKKKAG 180
DB 121 DLPPESQRHOMALLNFFPDEKAYSEESRRVRNRKSKSNEGADGPNKKNKKKAG 180
QY 181 PPGNGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 240
DB 181 PPGNGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 240
QY 241 AGTRENQPAVHLQGGSAIQVNDLSGGVNDWSRITMNPVKFLHPRSGELVLVDGT 300
DB 241 TGTRENQPAVHLQGGSAIQVNDLSGGVNDWSRITMNPVKFLHPRSGEL- 293
QY 301 YFIYSQVEVYINFTDFASVEVVDKPFLOCTRSIETGKTNTCTYTAGVCLLKARQKI 360
DB 294 -----EYIYNFTDFASVEVVDKPFLOCTRSIETGKTNTCTYTAGVCLLKARQKI 346
QY 361 AVKMVHADISINMSKHTTFFGAIRLGEAPAS 391
DB 347 AVKMVHADISINMSKHTTFFGAIRLGEAPAS 377

RESULT 5
US-09-949-016-11119
; Sequence 1119, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1119
; LENGTH: 260
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-11119
Query Match 66.9%; Score 1403; DB 2; Length 260;
Best Local Similarity 100.0%; Pred. No. 3.2e-110;
Matches 260; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 132 QWALLNFFPDEKPYSEESRRVRNRKSKSNEGADGPNKKNKKKAGPPGPPGPP 191
DB 1 QWALLNFFPDEKPYSEESRRVRNRKSKSNEGADGPNKKNKKKAGPPGPPGPP 60
QY 192 GPPGQGGPPGPIGPIGPIGPIGPIGPIGPIGPIGPIGPIGPIGPIGPIGPIG 251
DB 61 GPPGQGGPPGPIGPIGPIGPIGPIGPIGPIGPIGPIGPIGPIGPIGPIGPIG 120
QY 252 HLOGGSAIQVNDLSGGVNDWSRITMNPVKFLHPRSGELVLVDGTFTYISQVEVY 311
DB 121 HLOGGSAIQVNDLSGGVNDWSRITMNPVKFLHPRSGELVLVDGTFTYISQVEVY 180
QY 312 INFTDFASVEVVDKPFLOCTRSIETGKTNTCTYTAGVCLLKARQKI 371
DB 181 INFTDFASVEVVDKPFLOCTRSIETGKTNTCTYTAGVCLLKARQKI 240

QY 372 NMSKHTTFFGAIRLGEAPAS 391
DB 241 NMSKHTTFFGAIRLGEAPAS 260
RESULT 6
US-07-609-716-66
; Sequence 66, Application US/07609716
; Patent No. 5514581
; GENERAL INFORMATION:
; APPLICANT: Ferrari, Franco A.
; APPLICANT: Cappello, Joseph
; TITLE OF INVENTION: Functional Recombinantly Prepared
; TITLE OF INVENTION: Synthetic Protein Polymer
; NUMBER OF SEQUENCES: 118
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Flehr, Hohnbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: US
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/609,716
; FILING DATE: 06-NOV-1990
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Rowland, Bertram I
; REGISTRATION NUMBER: 20015
; REFERENCE/DOCKET NUMBER: A-55186-3/BI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-781-1889
; TELEFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 66:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 357 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-609-716-66
Query Match 14.0%; Score 294; DB 1; Length 357;
Best Local Similarity 32.4%; Pred. No. 1.2e-16;
Matches 69; Conservative 9; Mismatches 71; Indels 64; Gaps 3;
QY 71 ERGAESRLGSGTPTGTLSSLGGLDPSPTITSHLGQSPKQOPLPGEAALHSDSQ 128
DB 71 DRGDAGPKGADGSPGPGAPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPG 130
QY 129 DGHQWALLNFFPDEKPYSEESRRVRNRKSKSNEGADGPNKKNKKKAGPPGPP 188
DB 131 D-----RGDAGPKGADGSPGPGAPGPGPGPGPGPGPGPGPGPGPG 165
QY 189 GPPGPPGPPGPPGPIGPIGPIGPIGPIGPIGPIGPIGPIGPIGPIGPIG 212
DB 166 GPPGAPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP 225
QY 213 -VMGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP 244
DB 226 GPPGAPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP 258
RESULT 7
US-08-642-255-33
; Sequence 33, Application US/08642255
; Patent No. 5773249

```

; GENERAL INFORMATION:
; APPLICANT: CAPPELLO, Joseph
; APPLICANT: FERRARI, Franco A.
; TITLE OF INVENTION: High Molecular Weight Collagen-Like
; TITLE OF INVENTION: Protein Polymers
; NUMBER OF SEQUENCES: 135
; CORRESPONDENCE ADDRESS:
; ADDRESSES: FLEHR, HOBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/642,255
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: ROWLAND, Bertram I.
; REGISTRATION NUMBER: 20,015
; REFERENCE/DOCKET NUMBER: A55556-3/BIR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 494-8700
; TELEFAX: (415) 494-8771
; TELEX: 910 277299 FHT UR
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 357 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-642-255-33

Query Match 14.0%; Score 294; DB 1; Length 357;
Best Local Similarity 32.4%; Pred. No. 1.2e-16;
Matches 69; Conservative 9; Mismatches 71; Indels 64; Gaps 3;

QY 71 ERGASRLGGSTGTTSLSLGLDPS--ITSHLGQSPKQOPLPGEAALHSDSQ 128
Db 71 DRGDAGPKGADSGPGAGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPG 130
QY 129 DGHQALLNFFPPDEKPYSEESRRVRNRKSKSNEGADGPVKNKKKKKAGPPGNGPP 188
Db 131 D-----RGDAGPKGADSGPGAGPGPGPGPGPGPGPGPGPGPG 165
QY 189 GPPGPPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPG 212
Db 166 GPPGAPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPG 225
QY 213 -VMGPPGPPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPG 244
Db 226 GPPGAPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPG 258

RESULT 8
US-08-475-411A-66
; Sequence 66, Application US/08475411A
; Patent No. 6140072
; GENERAL INFORMATION:
; APPLICANT: Ferrari, Franco A.
; APPLICANT: Cappello, Joseph
; TITLE OF INVENTION: Functional Recombinantly Prepared
; TITLE OF INVENTION: Synthetic Protein Polymer
; NUMBER OF SEQUENCES: 119
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400

```

```

; CITY: San Francisco
; STATE: CA
; COUNTRY: US
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,411A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/609,716
; FILING DATE: 06-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/269,429
; FILING DATE: 09-NOV-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/114,618
; FILING DATE: 29-OCT-1987
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 06/927,258
; FILING DATE: 04-NOV-1986
; ATTORNEY/AGENT INFORMATION:
; NAME: Trecartin, Richard F.
; REGISTRATION NUMBER: 31,801
; REFERENCE/DOCKET NUMBER: A-55186-9/RFT/MTK
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-781-1989
; TELEFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 66:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 357 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-475-411A-66

Query Match 14.0%; Score 294; DB 2; Length 357;
Best Local Similarity 32.4%; Pred. No. 1.2e-16;
Matches 69; Conservative 9; Mismatches 71; Indels 64; Gaps 3;

QY 71 ERGASRLGGSTGTTSLSLGLDPS--ITSHLGQSPKQOPLPGEAALHSDSQ 128
Db 71 DRGDAGPKGADSGPGAGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPG 130
QY 129 DGHQALLNFFPPDEKPYSEESRRVRNRKSKSNEGADGPVKNKKKKKAGPPGNGPP 188
Db 131 D-----RGDAGPKGADSGPGAGPGPGPGPGPGPGPGPGPGPG 165
QY 189 GPPGPPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPG 212
Db 166 GPPGAPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPG 225
QY 213 -VMGPPGPPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPG 244
Db 226 GPPGAPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPGPG 258

RESULT 9
US-08-478-029A-66
; Sequence 66, Application US/08478029A
; Patent No. 6184348
; GENERAL INFORMATION:
; APPLICANT: Ferrari, Franco A.
; APPLICANT: Cappello, Joseph
; TITLE OF INVENTION: Functional Recombinantly Prepared
; TITLE OF INVENTION: Synthetic Protein Polymer
; NUMBER OF SEQUENCES: 119
; CORRESPONDENCE ADDRESS:

```

ADDRESS: Flehr, Hombach, Test, Albritton & Herbert
STREET: Four Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: CA
COUNTRY: US
ZIP: 94111

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

THIS PAGE BLANK (USPTO)


```
|||||
Db 361 AVKMVHADISINNSKHTTFFGAIRLGEAPAS 391

RESULT 2
US-09-729-658B-2
; Sequence 2, Application US/09729658B
; Publication No. US20030023991A1
; GENERAL INFORMATION:
; APPLICANT: Zonana et al.
; TITLE OF INVENTION: Hypohydrotic ectodermal dysplasia genes and proteins
; FILE REFERENCE: 55924
; CURRENT APPLICATION NUMBER: US/09/729, 658B
; PRIOR FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: 09/342,681
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: 60/092,279
; PRIOR FILING DATE: 1998-07-09
; PRIOR APPLICATION NUMBER: 60/112,366
; PRIOR FILING DATE: 1998-12-15
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 391
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-729-658B-2

Query Match 100.0%; Score 2098; DB 3; Length 391;
Best Local Similarity 100.0%; Pred. No. 1.6e-137;
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGYPEVERRELLPAAAPRERGSGCGCGAPARAGSGNSCLLPLGFFGLSLALHLLTLCC 60
Db 1 MGYPEVERRELLPAAAPRERGSGCGCGAPARAGSGNSCLLPLGFFGLSLALHLLTLCC 60
QY 61 YLELRSELRRERGAESRLGSGTGTGTLSSLGGLDPSPTITSHLGQSPKQPLEPGE 120
Db 61 YLELRSELRRERGAESRLGSGTGTGTLSSLGGLDPSPTITSHLGQSPKQPLEPGE 120
QY 121 AALHSDSQDHQWALLNFFFPDEKPYSEESRRVRNRKSKSNEGADGPVNKKKKKAG 180
Db 121 AALHSDSQDHQWALLNFFFPDEKPYSEESRRVRNRKSKSNEGADGPVNKKKKKAG 180
QY 181 PPGNPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 240
Db 181 PPGNPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 240
QY 241 AGTRENQPAVHLQGGSAIQVKNLSSGVLNDSRITWNPKVKLHPRSGELEVLVDGT 300
Db 241 AGTRENQPAVHLQGGSAIQVKNLSSGVLNDSRITWNPKVKLHPRSGELEVLVDGT 300
QY 301 YFIYSQVEVYIINFTDFASYEVVVDKPFLOCTRSIETGKTNTNTCYTAGVCLLKARQKI 360
Db 301 YFIYSQVEVYIINFTDFASYEVVVDKPFLOCTRSIETGKTNTNTCYTAGVCLLKARQKI 360
QY 361 AVKMVHADISINNSKHTTFFGAIRLGEAPAS 391
Db 361 AVKMVHADISINNSKHTTFFGAIRLGEAPAS 391

RESULT 3
US-10-218-547-42
; Sequence 42, Application US/10218547
; Publication No. US20030100074A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Methods And Compositions For Treating Metabolic Bone Diseases Rel
; FILE REFERENCE: PF561
; CURRENT APPLICATION NUMBER: US/10/218, 547
; CURRENT FILING DATE: 2002-08-15
; PRIOR APPLICATION NUMBER: 60/312,542
```

```
; PRIOR FILING DATE: 2001-08-16
; PRIOR APPLICATION NUMBER: 60/330,761
; PRIOR FILING DATE: 2001-10-30
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 42
; LENGTH: 391
; TYPE: PRT
; ORGANISM: human
US-10-218-547-42

Query Match 100.0%; Score 2098; DB 4; Length 391;
Best Local Similarity 100.0%; Pred. No. 1.6e-137;
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGYPEVERRELLPAAAPRERGSGCGCGAPARAGSGNSCLLPLGFFGLSLALHLLTLCC 60
Db 1 MGYPEVERRELLPAAAPRERGSGCGCGAPARAGSGNSCLLPLGFFGLSLALHLLTLCC 60
QY 61 YLELRSELRRERGAESRLGSGTGTGTLSSLGGLDPSPTITSHLGQSPKQPLEPGE 120
Db 61 YLELRSELRRERGAESRLGSGTGTGTLSSLGGLDPSPTITSHLGQSPKQPLEPGE 120
QY 121 AALHSDSQDHQWALLNFFFPDEKPYSEESRRVRNRKSKSNEGADGPVNKKKKKAG 180
Db 121 AALHSDSQDHQWALLNFFFPDEKPYSEESRRVRNRKSKSNEGADGPVNKKKKKAG 180
QY 181 PPGNPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 240
Db 181 PPGNPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 240
QY 241 AGTRENQPAVHLQGGSAIQVKNLSSGVLNDSRITWNPKVKLHPRSGELEVLVDGT 300
Db 241 AGTRENQPAVHLQGGSAIQVKNLSSGVLNDSRITWNPKVKLHPRSGELEVLVDGT 300
QY 301 YFIYSQVEVYIINFTDFASYEVVVDKPFLOCTRSIETGKTNTNTCYTAGVCLLKARQKI 360
Db 301 YFIYSQVEVYIINFTDFASYEVVVDKPFLOCTRSIETGKTNTNTCYTAGVCLLKARQKI 360
QY 361 AVKMVHADISINNSKHTTFFGAIRLGEAPAS 391
Db 361 AVKMVHADISINNSKHTTFFGAIRLGEAPAS 391

RESULT 4
US-10-012-605C-6
; Sequence 6, Application US/10012605C
; Publication No. US20030104602A1
; GENERAL INFORMATION:
; APPLICANT: Haldeman, Betty A.
; APPLICANT: Thayer, Edward C.
; APPLICANT: Sheppard, Paul O.
; TITLE OF INVENTION: ADIPOCYTE COMPLEMENT RELATED PROTEIN
; FILE REFERENCE: 00-111
; CURRENT APPLICATION NUMBER: US/10/012,605C
; CURRENT FILING DATE: 2002-08-14
; PRIOR APPLICATION NUMBER: US 60/254,019
; PRIOR FILING DATE: 2000-12-07
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 391
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-012-605C-6

Query Match 100.0%; Score 2098; DB 4; Length 391;
Best Local Similarity 100.0%; Pred. No. 1.6e-137;
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGYPEVERRELLPAAAPRERGSGCGCGAPARAGSGNSCLLPLGFFGLSLALHLLTLCC 60
```


Db 241 AGTRENQPAVHLQGGGSAIQVKNDSLGGVLDNWSRITWNPVKVFLHPRSGELEVLDGT 300
QY 301 YFIYSQVEVYINFTDPASVEVVVDEKPFLOCTRSIETGKTNNTCYTAGVCLLKARQKI 360
Db 301 YFIYSQVEVYINFTDPASVEVVVDEKPFLOCTRSIETGKTNNTCYTAGVCLLKARQKI 360
QY 361 AVQMVHADISINMSKHTTFFGAIRLGEAPAS 391
Db 361 AVQMVHADISINMSKHTTFFGAIRLGEAPAS 391

RESULT 7
US-10-719-205-6
; Sequence 6, Application US/10719205
; Publication No. US20040086971A1
; GENERAL INFORMATION:
; APPLICANT: Haldeman, Betty A.
; APPLICANT: Thayer, Edward C.
; APPLICANT: Sheppard, Paul O.
; TITLE OF INVENTION: ADIPOCYTE COMPLEMENT RELATED PROTEIN
; FILE OF INVENTION: ZACRP3X2
; FILE REFERENCE: 00-111
; CURRENT APPLICATION NUMBER: US/10/719,205
; CURRENT FILING DATE: 2003-11-21
; PRIOR APPLICATION NUMBER: US/10/012,605C
; PRIOR FILING DATE: 2002-08-14
; PRIOR APPLICATION NUMBER: US 60/254,019
; PRIOR FILING DATE: 2000-12-07
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 391
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-719-205-6

Query Match 100.0%; Score 2098; DB 4; Length 391;
Best Local Similarity 100.0%; Pred. No. 1.6e-137;
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGYPEVERRELLPAAAPRERGSGCGCGGAPAPAGEGNSCLLFLGFFGLSLALHLTLCC 60
Db 1 MGYPEVERRELLPAAAPRERGSGCGCGGAPAPAGEGNSCLLFLGFFGLSLALHLTLCC 60
QY 61 YLELRSELRRERGAESRLGGSGTPTGTSLSLGLDPSPTTSHLGQSPKQOQPLEPGE 120
Db 61 YLELRSELRRERGAESRLGGSGTPTGTSLSLGLDPSPTTSHLGQSPKQOQPLEPGE 120
QY 121 AALHSDSQDGHQWALLNFFFPDEKPYSEESRRVRNRKSKSNEGADGPVKNKKKKKAG 180
Db 121 AALHSDSQDGHQWALLNFFFPDEKPYSEESRRVRNRKSKSNEGADGPVKNKKKKKAG 180
QY 181 PPGNGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 240
Db 181 PPGNGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 240
QY 241 AGTRENQPAVHLQGGGSAIQVKNDSLGGVLDNWSRITWNPVKVFLHPRSGELEVLDGT 300
Db 241 AGTRENQPAVHLQGGGSAIQVKNDSLGGVLDNWSRITWNPVKVFLHPRSGELEVLDGT 300
QY 301 YFIYSQVEVYINFTDPASVEVVVDEKPFLOCTRSIETGKTNNTCYTAGVCLLKARQKI 360
Db 301 YFIYSQVEVYINFTDPASVEVVVDEKPFLOCTRSIETGKTNNTCYTAGVCLLKARQKI 360
QY 361 AVQMVHADISINMSKHTTFFGAIRLGEAPAS 391
Db 361 AVQMVHADISINMSKHTTFFGAIRLGEAPAS 391

RESULT 8
US-11-028-780-42
; Sequence 42, Application US/11028780
; Publication No. US20050163747A1

; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.,
; TITLE OF INVENTION: Heteromultimeric TNF Ligand Family members
; FILE REFERENCE: PFS59C1
; CURRENT APPLICATION NUMBER: US/11/028,780
; CURRENT FILING DATE: 2005-01-05
; PRIOR APPLICATION NUMBER: 10/202,062
; PRIOR FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,838
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 42
; LENGTH: 391
; TYPE: PRT
; ORGANISM: human
US-11-028-780-42
Query Match 100.0%; Score 2098; DB 6; Length 391;
Best Local Similarity 100.0%; Pred. No. 1.6e-137;
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGYPEVERRELLPAAAPRERGSGCGCGGAPAPAGEGNSCLLFLGFFGLSLALHLTLCC 60
Db 1 MGYPEVERRELLPAAAPRERGSGCGCGGAPAPAGEGNSCLLFLGFFGLSLALHLTLCC 60
QY 61 YLELRSELRRERGAESRLGGSGTPTGTSLSLGLDPSPTTSHLGQSPKQOQPLEPGE 120
Db 61 YLELRSELRRERGAESRLGGSGTPTGTSLSLGLDPSPTTSHLGQSPKQOQPLEPGE 120
QY 121 AALHSDSQDGHQWALLNFFFPDEKPYSEESRRVRNRKSKSNEGADGPVKNKKKKKAG 180
Db 121 AALHSDSQDGHQWALLNFFFPDEKPYSEESRRVRNRKSKSNEGADGPVKNKKKKKAG 180
QY 181 PPGNGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 240
Db 181 PPGNGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 240
QY 241 AGTRENQPAVHLQGGGSAIQVKNDSLGGVLDNWSRITWNPVKVFLHPRSGELEVLDGT 300
Db 241 AGTRENQPAVHLQGGGSAIQVKNDSLGGVLDNWSRITWNPVKVFLHPRSGELEVLDGT 300
QY 301 YFIYSQVEVYINFTDPASVEVVVDEKPFLOCTRSIETGKTNNTCYTAGVCLLKARQKI 360
Db 301 YFIYSQVEVYINFTDPASVEVVVDEKPFLOCTRSIETGKTNNTCYTAGVCLLKARQKI 360
QY 361 AVQMVHADISINMSKHTTFFGAIRLGEAPAS 391
Db 361 AVQMVHADISINMSKHTTFFGAIRLGEAPAS 391

RESULT 9
US-11-142-736-8
; Sequence 8, Application US/11142736
; Publication No. US20050227283A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: Novel Drosophila Tumor Necrosis Factor Class Molecule ("DmTNF") ar
; FILE REFERENCE: D0016 DIV1
; CURRENT APPLICATION NUMBER: US/11/142,736
; CURRENT FILING DATE: 2005-06-01
; PRIOR APPLICATION NUMBER: 60/190,816
; PRIOR FILING DATE: 2000-03-21
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 8
; LENGTH: 391
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-11-142-736-8
Query Match 100.0%; Score 2098; DB 6; Length 391;

```
Best Local Similarity 100.0%; Pred. No. 1.6e-137;
Matches 391; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGYPEVERRELLPAAAPRERGSCGCGGAPARAGEGNSCLLFLGFFGLSLALHLTLCC 60
DB 1 MGYPEVERRELLPAAAPRERGSCGCGGAPARAGEGNSCLLFLGFFGLSLALHLTLCC 60
QY 61 YLELSELRRERGAESRLGSGTPTGTLSSIGGLDPSDPTSHLGQPSKQOOLEPGE 120
DB 61 YLELSELRRERGAESRLGSGTPTGTLSSIGGLDPSDPTSHLGQPSKQOOLEPGE 120
QY 121 AALHSDSQDHQWALLNFFPDDEKPYSEESRRVRNRKRSKNEGADGPVKNKKKKKAG 180
DB 121 AALHSDSQDHQWALLNFFPDDEKPYSEESRRVRNRKRSKNEGADGPVKNKKKKKAG 180
QY 181 PPGNGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 240
DB 181 PPGNGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 240
QY 241 AGTRENQPAVHLQOGSAIQVNDLSGGVNDWSRITMNPVKFKLHPRSGEVLVDGT 300
DB 241 AGTRENQPAVHLQOGSAIQVNDLSGGVNDWSRITMNPVKFKLHPRSGEVLVDGT 300
QY 301 YFIYSQVEVYINFTDFASVEVVVDEKPFLOCTRSIETGKNTYNTCTAGVCLLKAROKI 360
DB 301 YFIYSQVEVYINFTDFASVEVVVDEKPFLOCTRSIETGKNTYNTCTAGVCLLKAROKI 360
QY 361 AVKMVHADISINNSKHTTFFGAIRLGEAPAS 391
DB 361 AVKMVHADISINNSKHTTFFGAIRLGEAPAS 391

RESULT 10
US-09-813-329-9
; Sequence 9, Application US/09813329
; Patent No. US20020012968A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Suibb Company
; TITLE OF INVENTION: Novel Drosophila Tumor Necrosis Factor Class Mole
; FILE REFERENCE: D0016.np
; CURRENT APPLICATION NUMBER: US/09/813,329
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/190,816
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 391
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-813-329-9

Query Match 93.7%; Score 1965; DB 3; Length 391;
Best Local Similarity 94.6%; Pred. No. 2.7e-128;
Matches 370; Conservative 1; Mismatches 20; Indels 0; Gaps 0;

QY 1 MGYPEVERRELLPAAAPRERGSCGCGGAPARAGEGNSCLLFLGFFGLSLALHLTLCC 60
DB 1 MGYPEVERRELLPAAAPRERGSCGCGGAPARAGEGNSCLLFLGFFGLSLALHLTLCC 60
QY 61 YLELSELRRERGAESRLGSGTPTGTLSSIGGLDPSDPTSHLGQPSKQOOLEPGE 120
DB 61 YLELSELRRERGAESRLGSGTPTGTLSSIGGLDPSDPTSHLGQPSKQOOLEPGE 120
QY 121 AALHSDSQDHQWALLNFFPDDEKPYSEESRRVRNRKRSKNEGADGPVKNKKKKKAG 180
DB 121 AALHSDSQDHQWALLNFFPDDEKPYSEESRRVRNRKRSKNEGADGPVKNKKKKKAG 180
QY 181 PPGNGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 240
DB 181 PPGNGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 240
QY 241 AGTRENQPAVHLQOGSAIQVNDLSGGVNDWSRITMNPVKFKLHPRSGEVLVDGT 300
DB 241 AGTRENQPAVHLQOGSAIQVNDLSGGVNDWSRITMNPVKFKLHPRSGEVLVDGT 300
QY 301 YFIYSQVEVYINFTDFASVEVVVDEKPFLOCTRSIETGKNTYNTCTAGVCLLKAROKI 360
DB 301 YFIYSQVEVYINFTDFASVEVVVDEKPFLOCTRSIETGKNTYNTCTAGVCLLKAROKI 360
QY 361 AVKMVHADISINNSKHTTFFGAIRLGEAPAS 391
DB 361 AVKMVHADISINNSKHTTFFGAIRLGEAPAS 391

RESULT 10
US-09-813-329-9
; Sequence 9, Application US/09813329
; Patent No. US20020012968A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Suibb Company
; TITLE OF INVENTION: Novel Drosophila Tumor Necrosis Factor Class Mole
; FILE REFERENCE: D0016.np
; CURRENT APPLICATION NUMBER: US/09/813,329
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/190,816
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 391
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-813-329-9

Query Match 93.7%; Score 1965; DB 3; Length 391;
Best Local Similarity 94.6%; Pred. No. 2.7e-128;
Matches 370; Conservative 1; Mismatches 20; Indels 0; Gaps 0;

QY 1 MGYPEVERRELLPAAAPRERGSCGCGGAPARAGEGNSCLLFLGFFGLSLALHLTLCC 60
DB 1 MGYPEVERRELLPAAAPRERGSCGCGGAPARAGEGNSCLLFLGFFGLSLALHLTLCC 60
QY 61 YLELSELRRERGAESRLGSGTPTGTLSSIGGLDPSDPTSHLGQPSKQOOLEPGE 120
DB 61 YLELSELRRERGAESRLGSGTPTGTLSSIGGLDPSDPTSHLGQPSKQOOLEPGE 120
QY 121 AALHSDSQDHQWALLNFFPDDEKPYSEESRRVRNRKRSKNEGADGPVKNKKKKKAG 180
DB 121 AALHSDSQDHQWALLNFFPDDEKPYSEESRRVRNRKRSKNEGADGPVKNKKKKKAG 180
QY 181 PPGNGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 240
DB 181 PPGNGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 240
```

```
QY 241 AGTRENQPAVHLQOGSAIQVNDLSGGVNDWSRITMNPVKFKLHPRSGEVLVDGT 300
DB 241 AGTRENQPAVHLQOGSAIQVNDLSGGVNDWSRITMNPVKFKLHPRSGEVLVDGT 300
QY 301 YFIYSQVEVYINFTDFASVEVVVDEKPFLOCTRSIETGKNTYNTCTAGVCLLKAROKI 360
DB 301 YFIYSQVEVYINFTDFASVEVVVDEKPFLOCTRSIETGKNTYNTCTAGVCLLKAROKI 360
QY 361 AVKMVHADISINNSKHTTFFGAIRLGEAPAS 391
DB 361 AVKMVHADISINNSKHTTFFGAIRLGEAPAS 391

RESULT 11
US-11-142-736-9
; Sequence 9, Application US/11142736
; Publication No. US20050227283A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Suibb Company
; TITLE OF INVENTION: Novel Drosophila Tumor Necrosis Factor Class Molecule ("DmTNF") ar
; FILE REFERENCE: D0016 DIV1
; CURRENT APPLICATION NUMBER: US/11/142,736
; PRIOR FILING DATE: 2005-06-01
; PRIOR APPLICATION NUMBER: 60/190,816
; PRIOR FILING DATE: 2000-03-21
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 391
; TYPE: PRT
; ORGANISM: Mus musculus
US-11-142-736-9

Query Match 93.7%; Score 1965; DB 6; Length 391;
Best Local Similarity 94.6%; Pred. No. 2.7e-128;
Matches 370; Conservative 1; Mismatches 20; Indels 0; Gaps 0;

QY 1 MGYPEVERRELLPAAAPRERGSCGCGGAPARAGEGNSCLLFLGFFGLSLALHLTLCC 60
DB 1 MGYPEVERRELLPAAAPRERGSCGCGGAPARAGEGNSCLLFLGFFGLSLALHLTLCC 60
QY 61 YLELSELRRERGAESRLGSGTPTGTLSSIGGLDPSDPTSHLGQPSKQOOLEPGE 120
DB 61 YLELSELRRERGAESRLGSGTPTGTLSSIGGLDPSDPTSHLGQPSKQOOLEPGE 120
QY 121 AALHSDSQDHQWALLNFFPDDEKPYSEESRRVRNRKRSKNEGADGPVKNKKKKKAG 180
DB 121 AALHSDSQDHQWALLNFFPDDEKPYSEESRRVRNRKRSKNEGADGPVKNKKKKKAG 180
QY 181 PPGNGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 240
DB 181 PPGNGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 240
QY 241 AGTRENQPAVHLQOGSAIQVNDLSGGVNDWSRITMNPVKFKLHPRSGEVLVDGT 300
DB 241 AGTRENQPAVHLQOGSAIQVNDLSGGVNDWSRITMNPVKFKLHPRSGEVLVDGT 300
QY 301 YFIYSQVEVYINFTDFASVEVVVDEKPFLOCTRSIETGKNTYNTCTAGVCLLKAROKI 360
DB 301 YFIYSQVEVYINFTDFASVEVVVDEKPFLOCTRSIETGKNTYNTCTAGVCLLKAROKI 360
QY 361 AVKMVHADISINNSKHTTFFGAIRLGEAPAS 391
DB 361 AVKMVHADISINNSKHTTFFGAIRLGEAPAS 391

RESULT 12
US-09-729-658B-4
; Sequence 4, Application US/09729658B
; Publication No. US20030023991A1
; GENERAL INFORMATION:
; APPLICANT: Zonana et al.
```

; TITLE OF INVENTION: Hypohydrotic ectodermal dysplasia genes and proteins
; FILE REFERENCE: 55924
; CURRENT APPLICATION NUMBER: US/09/729,658B
; CURRENT FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: 09/342,681
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: 60/092,279
; PRIOR FILING DATE: 1998-07-09
; PRIOR APPLICATION NUMBER: 60/112,366
; PRIOR FILING DATE: 1998-12-15
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 377
; TYPE: PRT
; ORGANISM: Mus musculus
;
US-09-729-658B-4
Query Match 89.3%; Score 1873; DB 3; Length 377;
Best Local Similarity 90.8%; Pred. No. 6.4e-122;
Matches 355; Conservative 2; Mismatches 20; Indels 14; Gaps 1;

QY 1 MGYPEVERRELLPAAAPRRGSGCGCGGAPAPRAGSGNSCLLFLGPFGLSLALHLTLCC 60
DB 1 MGYPEVERREPLPAAAPRRGSGCGCGRGAAPRAGEGNSCRFLGPFGLSLALHLTLCC 60
QY 61 YLELRSELRRGAESRLGSGTPTGTSGLTSSLGGLDPSPTSHLGGQSPKQOPLRGE 120
DB 61 YLELRSELRRGTERSLGGGAPGTSGLTSSPSGLDPVGPITRHLGGQSPFOOQPLRGE 120
QY 121 AALHSDSQDQHOMALLNFFPPDEKPYSEBESRRVRNRKRSKNEGADGPVNKKKGKAG 180
DB 121 DPLPPSQDQHOMALLNFFPPDEKAYSEBESRRVRNRKRSKSEGGADGPVNKKKGKAG 180
QY 181 PPGNPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 240
DB 181 PPGNPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 240
QY 241 AGTRENQPAVHLQGGGSAIQVNDLSGGVLDWSRITWNPKVKLHPRSGEVLVDGT 300
DB 241 TGTRENQPAVHLQGGGSAIQVNDLSGGVLDWSRITWNPKVKLHPRSGEL----- 293
QY 301 YFTYSQVEVYINFTDFASVEVVDKPFLOQTRSIETGKTNVNTCYTAGVCLLKARQKI 360
DB 294 -----EYVYINFTDFASVEVVDKPFLOQTRSIETGKTNVNTCYTAGVCLLKARQKI 346
QY 361 AVKMVHADISINMSKHTTFFGAIRLGEAPAS 391
DB 347 AVKMVHADISINMSKHTTFFGAIRLGEAPAS 377

RESULT 13
US-10-503-999-6
; Sequence 6, Application US/10503999
; Publication No. US20050152872A1
; GENERAL INFORMATION:
; APPLICANT: Gaide, Olivier
; APPLICANT: Schneider, Pascal
; APPLICANT: Tschopp, Jurg
; TITLE OF INVENTION: Fusion Constructs Containing Active Sections Of TNF Ligands
; FILE REFERENCE: 11436*13
; CURRENT APPLICATION NUMBER: US/10/503,999
; PRIOR FILING DATE: 2004-08-09
; PRIOR APPLICATION NUMBER: DE 102 05 368.5
; PRIOR FILING DATE: 2002-02-10
; PRIOR FILING DATE: 2002-02-11
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
; LENGTH: 412
; TYPE: PRT
; ORGANISM: artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
US-10-503-999-8
Query Match 36.2%; Score 759; DB 5; Length 410;
Best Local Similarity 86.4%; Pred. No. 1.7e-44;
Matches 153; Conservative 1; Mismatches 15; Indels 8; Gaps 3;

QY 215 GPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 274
DB 242 GKRSPQPK 295
QY 275 SRITWNPKVKLHPRSGEVLVDGTFTYFYSQVEVYINFTDFASVEVVDKPFLOCTR 334
DB 296 SRITWNPKVKLHPRSGEVLVDGTFTYFYSQ--VYINFTDFASVEVVDKPFLOCTR 353
QY 335 SIETGKTNVNTCYTAGVCLLKARQKIAMVNHADISINMSKHTTFFGAIRLGEAPAS 391
DB 354 SIETGKTNVNTCYTAGVCLLKARQKIAMVNHADISINMSKHTTFFGAIRLGEAPAS 410

RESULT 15
US-10-185-425-9
; Sequence 9, Application US/10185425
; Publication No. US20030053984A1
; GENERAL INFORMATION:

; TYPE: PRT
; ORGANISM: artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
US-10-503-999-6
Query Match 37.1%; Score 779; DB 5; Length 412;
Best Local Similarity 87.6%; Pred. No. 7.1e-46;
Matches 155; Conservative 1; Mismatches 15; Indels 6; Gaps 2;

QY 215 GPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 274
DB 242 GKRSPQPK 295
QY 275 SRITWNPKVKLHPRSGEVLVDGTFTYFYSQVEVYINFTDFASVEVVDKPFLOCTR 334
DB 296 SRITWNPKVKLHPRSGEVLVDGTFTYFYSQVEVYINFTDFASVEVVDKPFLOCTR 355
QY 335 SIETGKTNVNTCYTAGVCLLKARQKIAMVNHADISINMSKHTTFFGAIRLGEAPAS 391
DB 356 SIETGKTNVNTCYTAGVCLLKARQKIAMVNHADISINMSKHTTFFGAIRLGEAPAS 412

RESULT 14
US-10-503-999-8
; Sequence 8, Application US/10503999
; Publication No. US20050152872A1
; GENERAL INFORMATION:
; APPLICANT: Gaide, Olivier
; APPLICANT: Schneider, Pascal
; APPLICANT: Tschopp, Jurg
; TITLE OF INVENTION: Fusion Constructs Containing Active Sections Of TNF Ligands
; FILE REFERENCE: 11436*13
; CURRENT APPLICATION NUMBER: US/10/503,999
; PRIOR FILING DATE: 2004-08-09
; PRIOR APPLICATION NUMBER: PCT/EP02/09354
; PRIOR FILING DATE: 2002-08-21
; PRIOR APPLICATION NUMBER: DE 102 05 368.5
; PRIOR FILING DATE: 2002-02-10
; PRIOR APPLICATION NUMBER: DE 102 05 583.1
; PRIOR FILING DATE: 2002-02-11
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 8
; LENGTH: 410
; TYPE: PRT
; ORGANISM: artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
US-10-503-999-8
Query Match 36.2%; Score 759; DB 5; Length 410;
Best Local Similarity 86.4%; Pred. No. 1.7e-44;
Matches 153; Conservative 1; Mismatches 15; Indels 8; Gaps 3;

QY 215 GPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 274
DB 242 GKRSPQPK 295
QY 275 SRITWNPKVKLHPRSGEVLVDGTFTYFYSQVEVYINFTDFASVEVVDKPFLOCTR 334
DB 296 SRITWNPKVKLHPRSGEVLVDGTFTYFYSQ--VYINFTDFASVEVVDKPFLOCTR 353
QY 335 SIETGKTNVNTCYTAGVCLLKARQKIAMVNHADISINMSKHTTFFGAIRLGEAPAS 391
DB 354 SIETGKTNVNTCYTAGVCLLKARQKIAMVNHADISINMSKHTTFFGAIRLGEAPAS 410

RESULT 15
US-10-185-425-9
; Sequence 9, Application US/10185425
; Publication No. US20030053984A1
; GENERAL INFORMATION:

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.
OM protein - protein search, using sw model
Run on: June 2, 2006, 17:42:48 ; Search time 15 Seconds
(without alignments)
301.465 Million cell updates/sec

Title: US-09-729-658C-2
Perfect score: 2098
Sequence: 1 MGYPEVERRELLPAAAPRER.....NMSKHTTFGAIRLGEAPAS 391

Scoring table:
Gapop 10.0 , Gapext 0.5

Searched: 58871 seqs, 11565156 residues

Total number of hits satisfying chosen parameters: 58871

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA New.*
1: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US09_NEW_PUB pep.*
2: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US06_NEW_PUB pep.*
3: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US07_NEW_PUB pep.*
4: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US08_NEW_PUB pep.*
5: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/PCT_NEW_PUB pep.*
6: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US10_NEW_PUB pep.*
7: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US11_NEW_PUB pep.*
8: /EMC_Celerra_SIDS3/ptodata/2/pubpaa/US60_NEW_PUB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	265.5	12.7	1464	US-10-985-570-2	Sequence 2, Appli
2	263	12.5	1366	US-10-985-570-3	Sequence 3, Appli
3	256	12.2	1466	US-10-985-570-1	Sequence 1, Appli
4	252	12.0	744	US-10-505-928-493	Sequence 493, App
5	250	11.9	717	US-10-505-928-438	Sequence 438, App
6	244.5	11.7	314	US-11-293-697-3733	Sequence 3733, Ap
7	220.5	10.5	520	US-10-196-743-332	Sequence 332, App
8	186.5	8.9	399	US-10-953-349-22288	Sequence 22288, A
9	184	8.8	37	US-10-499-266-22	Sequence 22, Appl
10	150	7.1	542	US-11-293-697-4713	Sequence 4713, Ap
11	146	7.0	1085	US-10-505-928-343	Sequence 343, App
12	143	6.8	467	US-11-293-697-3606	Sequence 3606, Ap
13	140.5	6.7	411	US-11-293-697-4496	Sequence 4496, Ap
14	136.5	6.5	1460	US-11-247-437-14	Sequence 14, Appl
15	136	6.5	243	US-11-296-092-42	Sequence 42, Appl
16	133	6.3	262	US-10-953-349-17300	Sequence 17300, A
17	133	6.3	320	US-10-953-349-17299	Sequence 17299, A
18	133	6.3	502	US-10-511-937-3008	Sequence 3008, Ap
19	132	6.3	161	US-10-953-349-18169	Sequence 18169, A
20	132	6.3	167	US-10-953-349-18168	Sequence 18168, A
21	131.5	6.3	278	US-10-861-934-16	Sequence 16, Appl
22	131.5	6.3	278	US-10-861-934-26	Sequence 26, Appl
23	131.5	6.3	593	US-11-293-697-3668	Sequence 3668, Ap
24	131	6.2	250	US-11-296-092-76	Sequence 76, Appl
25	129.5	6.2	382	US-10-478-743B-2	Sequence 2, Appli

26	128.5	6.1	318	6	US-10-953-349-36014	Sequence 36014, A
27	128.5	6.1	410	6	US-10-953-349-36013	Sequence 36013, A
28	128.5	6.1	448	6	US-10-953-349-36012	Sequence 36012, A
29	128	6.1	172	6	US-10-953-349-18002	Sequence 18002, A
30	127	6.1	602	7	US-11-293-697-3416	Sequence 3416, Ap
31	126.5	6.0	416	7	US-11-293-697-4319	Sequence 4319, Ap
32	125.5	6.0	415	7	US-11-293-697-4534	Sequence 4534, Ap
33	125	6.0	953	7	US-11-312-958-56	Sequence 56, Appl
34	123.5	5.9	523	6	US-10-953-349-35575	Sequence 35575, A
35	123.5	5.9	783	6	US-10-511-937-2985	Sequence 2985, A
36	120.5	5.7	257	6	US-10-953-349-27984	Sequence 27984, A
37	120.5	5.7	265	6	US-10-953-349-27983	Sequence 27983, A
38	119.5	5.7	133	6	US-10-953-349-29375	Sequence 29375, A
39	119.5	5.7	213	6	US-10-953-349-29374	Sequence 29374, A
40	119.5	5.7	455	6	US-10-953-349-28615	Sequence 28615, A
41	119.5	5.7	472	6	US-10-953-349-28614	Sequence 28614, A
42	119.5	5.7	512	6	US-10-953-349-28613	Sequence 28613, A
43	119.5	5.7	663	7	US-11-293-697-2477	Sequence 2477, Ap
44	119	5.7	950	7	US-11-293-697-4492	Sequence 4492, Ap
45	118	5.6	281	6	US-10-511-937-2473	Sequence 2473, Ap

ALIGNMENTS

RESULT 1

US-10-985-570-2
; Sequence 2, Application US/10985570
; Publication No. US20060100138A1
; GENERAL INFORMATION:
; APPLICANT: Olsen, David R.
; APPLICANT: Polarek, James W.
; APPLICANT: Yang, Chunlin
; TITLE OF INVENTION: IMPLANTABLE COLLAGEN COMPOSITIONS
; FILE REFERENCE: FP0305 US
; CURRENT APPLICATION NUMBER: US/10/985,570
; CURRENT FILING DATE: 2004-11-10
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 2
; LENGTH: 1464
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-985-570-2

Query Match 12.7%; Score 265.5; DB 6; Length 1464;
Best Local Similarity 29.7%; Pred. No. 1.8e-10;
Matches 78; Conservative 16; Mismatches 80; Indels 89; Gaps 9;

QY	13	PAAAPRGRSQ-----CGCGAPARAGE-----GNSCLLFLGFFGLSLHLTLCCY 61
DB	774	PAGAPGDKGESGSPGAGPTGARGAPGDRGEPGCPA-----GFAG-----815
QY	62	LELRSELRRERGAESRLGSGTPTGTLSSLGGLDPSITSHLGOPSPKQOPLPGEA 121
DB	816	-----PPGADGQCAKGEFGDAGKADGADGPPGAGP-----AGPPFIGNVGAPGA- 861
QY	122	ALHSDSQDGHQALLNFFFPDEKPYSEBESRRVRNRKRSKNEGADGPVKNKKKGKAGP 181
DB	862	-----KGARGSAGPPGATGFPFGAAGRVGP 885
QY	182	PGFNGPFGPPGPPGP-----QGPPFIGIPGIPGTTVNGPPGPPGPPGPPGPPGPPG 235
DB	886	PGPSGNAGPPGPPGPPGAGKEGKGPRGTGAPRGPGE--VGPPGPPGPPGAGKSPGADGPA 943
QY	236	GAADKAGTFR--ENQPAVVHLOGO 256
DB	944	GAGTFFGQGIAGQGVVGLPCQ 966

RESULT 2

US-10-985-570-3
; Sequence 3, Application US/10985570

```
; Publication No. US20060100138A1
; GENERAL INFORMATION:
; APPLICANT: Olsen, David R.
; APPLICANT: Polarek, James W.
; APPLICANT: Yang, Chunlin
; TITLE OF INVENTION: IMPLANTABLE COLLAGEN COMPOSITIONS
; FILE REFERENCE: FP0305 US
; CURRENT APPLICATION NUMBER: US/10/985,570
; CURRENT FILING DATE: 2004-11-10
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 1366
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-985-570-3

Query Match      12.5%; Score 263; DB 6; Length 1366;
Best Local Similarity 32.9%; Pred. No. 2.5e-10;
Matches 80; Conservative 10; Mismatches 89; Indels 64; Gaps 10;

QY 13 PAAAPRERSQCGCG----GAPARAGE-GNSCLLFLGFFGLSLALHLTLCCYLELRSE 67
DB 638 PSGLPGERGAAGTGGKGBEPLRGEIGNP-----GRDG----- 673

QY 68 LRRERGAESRLGGSGTGTGTLSSLSGLDPSDPSITSHLQSPKQOPLPGEAALHSDS 127
DB 674 ---ARGAHGAVGAPGAGTAGDRGEAGAAGPAGP-----AGPRGSPGERGEVG----- 718

QY 128 QDGHQWALLNFFPDEKPYSEESRRVRNRKSKSNEGADGPVKNKKKAGKAGPPGNGPP 188
DB 719 -----PAGPNGFAGPAGAAQCPGAKGERGKGGKGENGVVGT-----GPVGAAGPAGPN 768

QY 186 GPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP 239
DB 769 GPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP 826

QY 240 KAG 242
DB 827 PVG 829

RESULT 3
US-10-985-570-1
; Sequence 1, Application US/10985570
; Publication No. US20060100138A1
; GENERAL INFORMATION:
; APPLICANT: Olsen, David R.
; APPLICANT: Polarek, James W.
; APPLICANT: Yang, Chunlin
; TITLE OF INVENTION: IMPLANTABLE COLLAGEN COMPOSITIONS
; FILE REFERENCE: FP0305 US
; CURRENT APPLICATION NUMBER: US/10/985,570
; CURRENT FILING DATE: 2004-11-10
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 1466
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-985-570-1

Query Match      12.2%; Score 256; DB 6; Length 1466;
Best Local Similarity 28.8%; Pred. No. 7.8e-10;
Matches 86; Conservative 12; Mismatches 97; Indels 104; Gaps 10;

QY 13 PAAAPRERSQCG----CGCGGAPARAGEGNSCLLFLGFFGLSLALHLTLCCYLELRSEL 68
DB 793 PRGSPGERGETGPPGAPGAPGAPGAPGAPGAPGAPGAPGAPGAPGAPGAPGAPGAP 823

QY 69 RRRERGAESRLGGSGTGTGTLSSLSGLDPSDPSITSHLQSPKQOPLPGEAALHSDS 128
DB 824 KGERGAPGKGGGPPGVAGTPCGGSPAGPPGP-----QGVKGERGSPGPGGAAG----- 873

QY 129 DGHQWALLNFFPDEKPYSEESRRVRNRKSKSNEGADGPVKNKKKAGKAGPPGNGPP 188
DB 874 -----FPG-----ARGLPGPNGNGNPPGPPGSGSPGKDGPPGAGNTGAP 914

QY 189 GPP-----GPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP 226
DB 915 GSPGVSGPKGDAGQPGKSGPAGQPPGAPGAPGLGIAGITGARGLAGAGPMPGPPGSPGP 974

QY 227 -----GPPGLOGPSGAADKAGT--RENQPAVVHLQSGSAIQVKND 265
DB 975 GVKSGSGKPGANGLSGERGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP 1033

RESULT 4
US-10-505-928-493
; Sequence 493, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; PRIOR FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 493
; LENGTH: 744
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-505-928-493

Query Match      12.0%; Score 252; DB 6; Length 744;
Best Local Similarity 25.6%; Pred. No. 7.3e-10;
Matches 112; Conservative 26; Mismatches 145; Indels 154; Gaps 17;

QY 2 GYVEVERRELLPAAAPRERSGOGCGGAPARAGSGNSCLLFLGFFGLSLALHLTLCCY 61
DB 328 GFPGGKGEQGLP-GLPGPPGLPGIKPGPPGPKGD-----RGMGGVPGAL----- 371

QY 62 LELRSELRERGAESRLGGSGTGTGTLSSLSGLDPSDPSITSHLQSPK-----Q 113
DB 372 -----GPRGKGPICAGGIGGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP 422

QY 114 QPLEP-GEAALHSDSQDGHQWALLNFFPDEKPYSEESRRVRNRKSKSNEGADGPVKN 172
DB 423 GPPGPKGBFGLQG-----PPGKPFILGEVGPQGM-----GLPGPIGP 460

QY 173 K-KGKKKA-----GPP-----GPNPDPGPPGPPGPPGPPGPPGPPGPPGPPGPP 199
DB 461 KGEAGKGVGPPGLVGPVGLLGPKEGPIGDQGLQPPGPIGIGGSGPIGPIGPIGPKGE 520

QY 200 PGIPGIPGIPG-----TTVMGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP 253
DB 521 PGLPGPPGPPGIPGKPGVAGLHGPPCKPGALGPGQGLGPPGPPGPPGPPGPPGPPGPP 580

QY 254 QGQ-----GSAIQVKNDLSGGVLNDSRITMN-----PKVF----- 284
DB 581 QGEYLPDMGLGIDGVKPPHAYGAKKNGGPPAYEMPAFTAELTAPFPVPVGAIVKFKLLY 640

QY 285 -----KLHPSRGEVLVDGTYFIYSQV-----EVYINFTDFA 318
DB 641 NGRQNYNPOTGIFTCEVPGVYFAFVHCKGGNVWVAFKQNEPVMYTYDEYKKGFLDQA 700

QY 319 SYEVVVDEKP-----ELQ 331
DB 701 SGSAVILLRPGDRVFLQ 717

RESULT 5
US-10-505-928-438
```


<p> ; SOFTWARE: PatentIn 3.2 ; SEQ ID NO 343 ; LENGTH: 1085 ; TYPE: PRT ; ORGANISM: Homo sapiens US-10-505-928-343 </p> <p> Query Match 7.0%; Score 146; DB 6; Length 1085; Best Local Similarity 30.7%; Pred. No. 0.0097; Matches 70; Conservative 32; Mismatches 70; Indels 56; Gaps 19; </p> <p> QY 113 QOPLGEGAAHSDQDGHQ-MALLNFFPDEKPYSEESRRVRNRKSKNEGADGPVK 171 DB 470 QOKLEKERECDAKTQKEMMOTLN-----KKKEKLEKETEHEKQVK-QQVADLTAQ 521 </p> <p> QY 172 NKKKGKKA-----GP-PG-PNGP-----PGPPGPPGPGQ--PGIPGIPGIPG 210 DB 522 LHELRSRAVCASIPGSPGAGCGPPSSVPGSLLPLPPPPPLPGMLPPPPPLP--PG 579 </p> <p> QY 211 TTVMGPPGPPGPPGPGQ-----PPGLOGPSGAADKAGTRENQPAVVH-LQGGGSAIQVND 265 DB 580 -----GPPPPGPPG--PLGAIMPPG--APMGLALK--KKSIPQPTNALKSFWNSKLPENK 629 </p> <p> QY 266 LSGGVLDNSRTWMPKVKFLHPRSGEVLVDGTYFYISQVVEYYIN 313 DB 630 LEGTV---WTEID-DTKVKFIL-----DLEDL-ERTFSAYORQDFFVN 668 </p> <p> RESULT 12 US-11-293-697-3606 ; Sequence 3606, Application US/11293697 ; Publication No. US20060105376A1 ; GENERAL INFORMATION: ; APPLICANT: HELIX RESEARCH INSTITUTE ; TITLE OF INVENTION: Novel full length cdna ; FILE REFERENCE: HI-A0106 ; CURRENT APPLICATION NUMBER: US/11/293,697 ; CURRENT FILING DATE: 2005-12-05 ; PRIOR APPLICATION NUMBER: US/10/108,260 ; PRIOR FILING DATE: 2002-03-28 ; NUMBER OF SEQ ID NOS: 5458 ; SOFTWARE: PatentIn Ver. 2.1 ; SEQ ID NO 3606 ; LENGTH: 467 ; TYPE: PRT ; ORGANISM: Homo sapiens US-11-293-697-3606 </p> <p> Query Match 6.8%; Score 143; DB 7; Length 467; Best Local Similarity 27.5%; Pred. No. 0.0067; Matches 63; Conservative 18; Mismatches 59; Indels 90; Gaps 13; </p> <p> QY 65 RSELRRGAESRLGSGTPTGTSGLSLGLDPPDPSITSHLGQPSKQOPLGEGAAH 124 DB 150 REQLEWER--ERRISSAAAPAS-----VETPLNSVLGDSAS-----EPG---LQ 189 </p> <p> QY 125 SDSQDGHQWALLNFFPDEKPYSEESRRVRNRKSKNEGADGPVKKKKKGKAGPPG 184 DB 190 AASQ-----PAETP-----SQQGVILGFLAP 210 </p> <p> QY 185 NGPPGPPGPPGPPG-----QGPPGIPGIPGIPGTTVMGPPGPPGPPG-----PGPPGLQ 232 DB 211 --PPPPPLPGQAASVALPPPPGPPPPPLPST---GPPPPPPPPPLPNQVPPPPPPPP 265 </p> <p> QY 233 GP-----SGAADKAGTRENQPAVVH-LQGGGSAIQVNDLSGGVLDNSRI 277 DB 266 APPLPASGFFLASMSDNRP-----LTGLAAAI-----AGAKLRKVS RM 304 </p> <p> RESULT 13 US-11-293-697-4496 ; Sequence 4496, Application US/11293697 ; Publication No. US20060105376A1 </p>	<p> ; GENERAL INFORMATION: ; APPLICANT: HELIX RESEARCH INSTITUTE ; TITLE OF INVENTION: Novel full length cdna ; FILE REFERENCE: H1-A0106 ; CURRENT APPLICATION NUMBER: US/11/293,697 ; CURRENT FILING DATE: 2005-12-05 ; PRIOR APPLICATION NUMBER: US/10/108,260 ; PRIOR FILING DATE: 2002-03-28 ; NUMBER OF SEQ ID NOS: 5458 ; SOFTWARE: PatentIn Ver. 2.1 ; SEQ ID NO 4496 ; LENGTH: 411 ; TYPE: PRT ; ORGANISM: Homo sapiens US-11-293-697-4496 </p> <p> Query Match 6.7%; Score 140.5; DB 7; Length 411; Best Local Similarity 29.4%; Pred. No. 0.0087; Matches 92; Conservative 14; Mismatches 94; Indels 113; Gaps 26; </p> <p> QY 11 LLPAA---APRERGSGCGGAPARAGEGNSCLLFLGFFGLSLALHLLTLCCYLELRSE 67 DB 46 LVPAAASAP-QRGPSSI-----LPAAEPTICSLHQAWGG-----PG 81 </p> <p> QY 68 LRRERGAESRLG-GSG-----TPGTSGLSLGLDPPDPSITSHLGQPSKQOPLGEGAA 121 DB 82 CRAQKGIPAAALSPGPIAPIPGPAQIPGPIGPGSIP-GPIGPAQIPSPAPPI-EGPI 139 </p> <p> QY 122 ALHSDSQDGHQWALLNFFPDEKPYSEESRRVRNRKSKNEGADGPVKKKKKGKAGP 181 DB 140 S-----GPVQIPGPIGPIGPIIS-----GPAPIPGIPGPIGPIGPIG 178 </p> <p> QY 182 -PGPN-GP-PGP-----PGP-----QGP-PG-IPG-IPG----- 207 DB 179 IFGNPPIGPIGPIGPIGPIGPIGPIGPIGPIGPIGPIGPIGPIGPIGPIGPIGPIG 238 </p> <p> QY 208 IPTGTVMG-----PGP-PGP-----PGP-----QGPLOGPSGAADKAGTRE-- 245 DB 239 IFG-PIGPNFGLIFGPIGPIGPIGPIGPIGPIGPIGPIGPIGPIGPIGPIGPIG 297 </p> <p> QY 246 NOPAVVHLQGGGS 258 DB 298 SLFGLALILGPGS 310 </p> <p> RESULT 14 US-11-247-437-14 ; Sequence 14, Application US/11247437 ; Publication No. US20060110753A1 ; GENERAL INFORMATION: ; APPLICANT: Genzyme Corporation ; APPLICANT: Roberts, Bruce ; TITLE OF INVENTION: METHODS TO DIAGNOSE AND TREAT LUNG CANCER ; FILE REFERENCE: 5257C ; CURRENT APPLICATION NUMBER: US/11/247,437 ; CURRENT FILING DATE: 2005-10-11 ; PRIOR APPLICATION NUMBER: PCT/US2004/011193 ; PRIOR FILING DATE: 2004-04-12 ; PRIOR APPLICATION NUMBER: US 60/462,028 ; PRIOR FILING DATE: 2003-04-10 ; NUMBER OF SEQ ID NOS: 14 ; SOFTWARE: PatentIn version 3.1 ; SEQ ID NO 14 ; LENGTH: 1460 ; TYPE: PRT ; ORGANISM: Homo sapiens US-11-247-437-14 </p> <p> Query Match 6.5%; Score 136.5; DB 7; Length 1460; Best Local Similarity 24.5%; Pred. No. 0.054; Matches 75; Conservative 21; Mismatches 127; Indels 83; Gaps 13; </p> <p> QY 4 PEVERRELLPA-----AAPR-ERGSQCGCGGAPARAG----- 35 </p>
---	---

